

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
ΤΜΗΜΑ ΟΙΚΟΝΟΜΙΚΗΣ ΕΠΙΣΤΗΜΗΣ



ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
ΣΤΗΝ ΟΙΚΟΝΟΜΙΚΗ ΚΑΙ ΕΠΙΧΕΙΡΗΣΙΑΚΗ
ΣΤΡΑΤΗΓΙΚΗ

ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΚΑΙ ΒΡΑΧΥΠΡΟΘΕΣΜΗ
ΠΡΟΒΛΕΨΗ ΤΩΝ ΤΙΜΩΝ ΧΡΟΝΟΣΕΙΡΑΣ
ΚΡΥΠΤΟΝΟΜΙΣΜΑΤΩΝ, ΧΡΗΣΙΜΟΠΟΙΩΝΤΑΣ
ΤΕΧΝΙΚΕΣ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ

ΔΗΜΗΤΡΙΟΥ ΜΑΡΙΑ

Διπλωματική Εργασία υποβληθείσα στο Τμήμα Οικονομικών Επιστημών του Πανεπιστημίου Πειραιώς ως μέρος των απαιτήσεων για την απόκτηση Μεταπτυχιακού Διπλώματος Ειδίκευσης στην Οικονομική και Επιχειρησιακή Στρατηγική

Πειραιάς, Νοέμβριος 2022

UNIVERSITY OF PIRAEUS
DEPARTMENT OF ECONOMICS



MASTER PROGRAM IN
ECONOMIC AND BUSINESS STRATEGY

MODELING AND SHORT TERM FORECASTING OF
CRYPTOCURRENCY TIME SERIES PRICES USING
MACHINE LEARNING TECHNIQUES

By

DIMITRIOU MARIA

Master Thesis submitted to the Department of Economics of the University of Piraeus in partial fulfillment of the requirements for the degree of Master of Arts in Economic and Business Strategy

Piraeus, Greece, November 2022

Ευχαριστίες

Με την ολοκλήρωση της μεταπτυχιακής διπλωματικής μου εργασίας, θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες σε όλους όσους συνέβαλλαν στην εκπόνησή της.

Η ολοκλήρωση της μεταπτυχιακής αυτής εργασίας θα ήταν αδύνατη χωρίς τη πολύτιμη υποστήριξη της καθηγήτριας μου. Επ. Καθηγήτρια Κας Ιωάννας Κοκορέ και την ευχαριστώ πολύ για αυτό.

Επίσης, ευχαριστώ τον καθηγητή, κύριο Γεώργιο Καραγιαννίδη, Καθηγητή Ψηφιακών Τηλεπικοινωνιακών Συστημάτων του ΑΠΘ, ο οποίος με παρότρυνε να ακολουθήσω το συγκεκριμένο μεταπτυχιακό.

Καθώς επίσης τη συμφοιτήτρια μου Κόκκορη Μαρία για τη συμπαράσταση που μου παρείχε όλο αυτό το διάστημα.

Και τέλος, θα ήθελα εκφράσω την ευγνωμοσύνη μου στην οικογένειά μου για όλη τη στήριξη, τη συμπαράσταση και την κατανόησή τους, καθ' όλη τη διάρκεια των σπουδών μου.

ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΚΑΙ ΒΡΑΧΥΠΡΟΘΕΣΜΗ ΠΡΟΒΛΕΨΗ ΤΩΝ ΤΙΜΩΝ ΧΡΟΝΟΣΕΙΡΑΣ ΚΡΥΠΤΟΝΟΜΙΣΜΑΤΩΝ, ΧΡΗΣΙΜΟΠΟΙΩΝΤΑΣ ΤΕΧΝΙΚΕΣ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ

Σημαντικοί Όροι: Κρυπτονομίσματα, μοντέλα πρόβλεψης, ανάλυση χρονοσειράς, μηχανική μάθηση

Περίληψη

Λόγω της αυξανόμενης δημοτικότητάς τους και της εμπορικής αποδοχής τους, τα κρυπτονομίσματα διαδραματίζουν ολοένα και πιο ουσιαστικό ρόλο στην αλλαγή του χρηματοπιστωτικού συστήματος. Ενώ πολλοί άνθρωποι επενδύουν σε κρυπτονομίσματα, τα δυναμικά χαρακτηριστικά και η προβλεψιμότητα τους εξακολουθούν να είναι σε μεγάλο βαθμό άγνωστα, θέτοντας σε κίνδυνο τις επενδύσεις. Η μηχανική μάθηση καλείται να αντιμετωπίσει το πρόβλημα πρόβλεψης αυτών των νομισμάτων. Πολλές έρευνες προσπαθούν να εντοπίσουν το καλύτερο μοντέλο για την πρόβλεψη της τιμής του Bitcoin, ως το ισχυρότερο, αφήνοντας ένα σημαντικό ερευνητικό κενό στην πρόβλεψη και των υπολοίπων κρυπτονομισμάτων. Μία ολοκληρωμένη έρευνα θα εντοπίσει το καλύτερο δυνατό μοντέλο εφαρμογής στην χρονοσειρά τιμών και θα εξάγει σημαντικά συμπεράσματα στον τομέα των προβλέψεων.

Σε αυτή την εργασία, επιχειρούμε να προβλέψουμε την τιμή τεσσάρων κρυπτονομισμάτων (Bitcoin, Ethereum, Ripple και Tether) λαμβάνοντας υπόψη ποικιλία παραγόντων που επηρεάζουν την αξία τους με την υψηλότερη δυνατή ακρίβεια. Σαν μέθοδοι πρόβλεψης επιλέχθηκαν τρία διαφορετικής λογικής μοντέλα. Με την βοήθεια της Python και των βιβλιοθηκών της εφαρμόστηκαν στα δεδομένα των χρονοσειρών το παραδοσιακό μοντέλο αυτοπαλίνδρομου ολοκληρωμένου κινούμενου μέσου όρου (ARIMA), το επαναλαμβανόμενο νευρωνικό δίκτυο βραχυπρόθεσμης και μακροπρόθεσμης μνήμης (LSTM) και το μοντέλο δομής δένδρου αποφάσεων τυχαίου δάσους (Random Forest). Οι προβλέψεις δοκιμάστηκαν στα πραγματικά δεδομένα της χρονοσειράς και εφαρμόστηκε ο δείκτης ρίζας του μέσου τετραγωνικού σφάλματος (RMSE) για την αξιολόγηση και σύγκριση των μοντέλων.

«Μπορεί το όνειρο των ερευνητών και επενδυτών για πρόβλεψη της τιμής των κρυπτονομισμάτων να γίνει ένα ρεαλιστικό σενάριο;». Ενώ η παραδοσιακή μέθοδος Arima και

η μέθοδος δένδρων απόφασης Random Forest δεν παρείχαν ικανοποιητικές προβλέψεις, το μοντέλο LSTM εμφάνισε ελπιδοφόρα αποτελέσματα. Το μοντέλο LSTM μπορεί να διατηρήσει προηγούμενες πληροφορίες, οι οποίες μπορούν να βοηθήσουν σημαντικά στη βελτίωση της ικανότητας του μοντέλου να μαθαίνει ακολουθίες σημάτων και εγγενή μη γραμμικά μοτίβα.

MODELING AND SHORT TERM FORECASTING OF CRYPTOCURRENCY TIME SERIES PRICES USING MACHINE LEARNING TECHNIQUES

Keywords: Cryptocurrencies, Predictive models, Time series analysis, Machine Learning

Abstract

Due to their growing popularity and commercial acceptance, cryptocurrencies are playing an increasingly essential role in changing the financial system. While many people invest in cryptocurrency, the dynamic characteristics and predictability of them are still largely unknown, putting investments at risk. Machine learning is called upon to tackle the problem of predicting these currencies. Many researches are trying to identify the best model for predicting the price of Bitcoin, as the strongest one, leaving a significant research gap in the prediction of other cryptocurrencies as well. A comprehensive investigation will identify the best possible model with remarkable results in the forecast field.

In this paper, we attempt to predict the price of four cryptocurrencies (Bitcoin, Ethereum, Ripple and Tether) taking into account a variety of factors that affect their value with the highest possible accuracy. Three models of different logic were chosen as prediction methods. With the help of Python and its libraries, the traditional Autoregressive Integrated Moving Average (ARIMA) model, the Long Short-Term Memory model (LSTM) and the decision tree structure model Random Forest (RF) were applied to the time series data. The predictions were tested on the real time series data and the root mean square error (RMSE) was applied to evaluate and compare the models.

"Can researchers and investors' dream of cryptocurrency price prediction become a realistic scenario?". While the traditional Arima method and the RF decision tree method did not provide satisfactory predictions, the LSTM model showed promising results. The LSTM model can retain prior information, which can greatly help improve the model's ability to learn signal sequences and inherent non-linear patterns

Πίνακας περιεχομένων

1	Εισαγωγή	1
2	Ανασκόπηση βιβλιογραφίας	3
3	Εισαγωγή στα κρυπτονομίσματα	5
3.1	Εισαγωγή	5
3.2	Κρυπτονομίσματα και blockchain	5
3.3	Ιστορική αναδρομή.....	7
3.4	Δημοφιλέστερα κρυπτονομίσματα και blockchain.....	11
3.4.1	Κορυφαία Δίκτυα Blockchain	12
3.4.2	Κορυφαία Κρυπτονομίσματα.....	14
3.5	Πλεονεκτήματα και μειονεκτήματα των κρυπτονομισμάτων.....	17
3.6	Ομοιότητες και διαφορές κρυπτονομισμάτων και ηλεκτρονικού χρήματος.....	20
3.7	Ανακεφαλαίωση	21
Θεωρητικό υπόβαθρο μοντέλων πρόβλεψης		23
4.1	Εισαγωγή	23
4.2	Βασικά χαρακτηριστικά και ανάλυση χρονοσειρών	23
4.3	Ορισμός πρόβλεψης χρονοσειράς.....	30
4.4	Βήματα πρόβλεψης.....	30
4.5	Βασικοί στατιστικοί έλεγχοι	32
4.6	Κριτήρια αξιολόγησης της πρόβλεψης.....	33
4.7	Παρουσίαση των προς μελέτη μοντέλων πρόβλεψης	35
4.7.1	Auto-Regressive Integrated Moving Average (ARIMA)	36
4.7.2	Long Short-Term Memory (LSTM).....	38
4.7.3	Random Forest (RF).....	39
4.8	Ανακεφαλαίωση	40
5	Εφαρμογή των μοντέλων πρόβλεψης.....	41
5.1	Εισαγωγή	41
5.2	Πρόβλεψη της τιμής του BTC	42
5.2.1	Μέθοδος ARIMA.....	42
5.2.2	Μέθοδος LSTM	54
5.2.3	Μέθοδος Random Forest	58
5.3	Πρόβλεψη τιμής Ethereum	61
5.3.1	Μέθοδος ARIMA.....	61

5.3.2	Μέθοδος LSTM.....	65
5.3.3	Μέθοδος Random Forest	66
5.4	Πρόβλεψη της τιμής Ripple (XRP).....	68
5.4.1	Μέθοδος ARIMA.....	68
5.4.2	Μέθοδος LSTM	71
5.4.3	Μέθοδος Random Forest	72
5.5	Πρόβλεψη της τιμής Tether (USDT)	73
5.5.1	Μέθοδος ARIMA.....	74
5.5.2	Μέθοδος LSTM.....	77
5.5.3	Μέθοδος Random Forest	78
5.6	Συγκέντρωση τιμών – Στατιστικός Έλεγχος.....	79
5.7	Ανακεφαλαίωση	82
6	Συμπεράσματα	85

Κατάλογος Πινάκων

Πίνακας 5.1: Παρουσίαση τεσσάρων πρώτων ενδεικτικών τιμών του dataframe με αναπαράσταση της ώρας και της τιμής του BTC.....	42
Πίνακας 5.2: Αποτελέσματα ADF και KPSS test μετά την αλλαγή της μεταβλητότητας.....	44
Πίνακας 5.3: Αποτελέσματα ελέγχου ADF και KPSS μετά τις διαφοροποιήσεις.....	45
Πίνακας 5.4: Διαχωρισμός δεδομένων στη μέση, εύρεση μέσης τιμές και τυπικών αποκλίσεων κάθε μισού.....	49
Πίνακας 5.5: Αποτελέσματα Ljung-Box στα δεδομένα της χρονοσειράς.....	50
Πίνακας 5.6: Αποτελέσματα Ljung-Box στα υπόλοιπα της χρονοσειράς.....	50
Πίνακας 5.7: Χρήση Auto- Arima για σύγκριση και επιλογή του βέλτιστου μοντέλου.....	51
Πίνακας 5.8: Βασικό μοντέλο LSTM – πρόβλεψη BTC.....	56
Πίνακας 5.9: Βασικό μοντέλο LSTM - πρόβλεψη ETH.....	66
Πίνακας 5.10: Βασικό μοντέλο LSTM - πρόβλεψη XRP.....	71
Πίνακας 5.11: : Βασικό μοντέλο LSTM - πρόβλεψη USDT.....	77
Πίνακας 5.12: Τιμές RMSE κρυπτονομισμάτων για τα τρία μοντέλα μελέτης.....	80

Κατάλογος Εικόνων

Εικόνα 3.1: Αξία των δημοφιλέστερων δικτύων blockchain της αγοράς, Πηγή (6/10.2022): https://blockchain.coinmarketcap.com/?utm_source=coinmarketcap&utm_content=footer ..	13
Εικόνα 3.2: Τα 5 πιο κορυφαία κρυπτονομίσματα ανά κεφαλαιοποίηση αγοράς. Πηγή (6/10/2022): https://coinmarketcap.com/	14
Εικόνα 4.1: Η δομή ενός αλγόριθμου μακράς βραχυπρόθεσμης μνήμης (LSTM).....	38
Εικόνα 5.1: Παρουσίαση ιστοσελίδας εξόρυξης δεδομένων χρονοσειράς τιμών κρυπτονομισμάτων (Στιγμιότυπο οθόνης 2/10/2022).	41
Εικόνα 5.2: Χρήση της συνάρτησης grid search forecaster για την εύρεση καλύτερων υπερπαραμέτρων.	60

Κατάλογος Διαγραμμάτων

Διάγραμμα 4.1: Παράδειγμα στασιμότητας και μη σε μία χρονοσειρά. Πηγή: https://blogs.cisco.com/analytics-automation/arima2	24
Διάγραμμα 4.2: Παράδειγμα χρονοσειράς με τάση. Πηγή: https://www.aptech.com/blog/introduction-to-the-fundamentals-of-time-series-data-and-analysis/	25
Διάγραμμα 4.3: Παράδειγμα χρονοσειράς με εποχικότητα. Πηγή: https://www.aptech.com/blog/introduction-to-the-fundamentals-of-time-series-data-and-analysis/	26
Διάγραμμα 4.4: Παράδειγμα χρονοσειράς με κυκλικότητα. Πηγή: https://robjhyndman.com/hyndsight/cyclights/	27
Διάγραμμα 4.5: Παράδειγμα ακραίας/ ασυνήθιστης τιμής σε χρονοσειρά. Πηγή: https://pro.arcgis.com/en/pro-app/latest/tool-reference/space-time-pattern-mining/understanding-outliers-in-time-series-analysis.htm	28
Διάγραμμα 5.1: Αναπαράσταση ιστορικών στοιχείων χρονοσειράς για την τιμή του BTC. ...	43
Διάγραμμα 5.2: Εντοπισμός και αφαίρεση τάσης στην χρονοσειρά BTC.	44
Διάγραμμα 5.3: Οπτικοποίηση δεδομένων χρονοσειράς μετά από διαφοροποίηση για επίτευξη σταθερότητας.	46
Διάγραμμα 5.4: Έλεγχος τυχαιότητας δεδομένων σε διάγραμμα ACF.	47
Διάγραμμα 5.5: Έλεγχος τυχαιότητας δεδομένων σε διάγραμμα PACF.	47
Διάγραμμα 5.6: Κατανομές τιμών των δύο μισών της χρονοσειράς.	49
Διάγραμμα 5.7: Τεστ του μοντέλου Πρόβλεψης της τιμής του BTC- μέθοδος ARIMA.	52
Διάγραμμα 5.8: Τεστ του βελτιωμένου μοντέλου Πρόβλεψης της τιμής του BTC- μέθοδος ARIMA.	53
Διάγραμμα 5.9: Πρόβλεψη της τιμής BTC για μία εβδομάδα (27/09/2022 - 4/10/2022) με διάστημα εμπιστοσύνης 95%.	54
Διάγραμμα 5.10: Πρόβλεψη της τιμής BTC για τρεις μήνες (27/09/2022 - 27/12/2022) με διάστημα εμπιστοσύνης 95%.	54
Διάγραμμα 5.11: Εφαρμογή του επιλεγμένου μοντέλου LSTM στα δεδομένα της χρονοσειράς.	56
Διάγραμμα 5.12: Τεστ του μοντέλου Πρόβλεψης της τιμής του BTC- μέθοδος LSTM.	57
Διάγραμμα 5.13: Μέρες εκπαίδευσης και δοκιμής του μοντέλου Random Forest για πρόβλεψη BTC.	58
Διάγραμμα 5.14: Δοκιμή προβλέψεων τιμής BTC με το μοντέλο Random Forest.	59
Διάγραμμα 5.15: Εύρεση βέλτιστου μοντέλου προβλέψεων τιμής BTC- Random Forest.	60
Διάγραμμα 5.16: : Αναπαράσταση ιστορικών στοιχείων χρονοσειράς για την τιμή του Ethereum.	61
Διάγραμμα 5.17: Οπτικοποίηση δεδομένων χρονοσειράς μετά από διαφοροποίηση για επίτευξη σταθερότητας.	62
Διάγραμμα 5.18: : Έλεγχος τυχαιότητας δεδομένων σε διάγραμμα ACF και PACF.	62
Διάγραμμα 5.19: Τεστ του μοντέλου Πρόβλεψης της τιμής του ETH- μέθοδος ARIMA.	64
Διάγραμμα 5.20: Πρόβλεψη της τιμής Ethereum για μία εβδομάδα (27/09/2022 - 4/10/2022) με διάστημα εμπιστοσύνης 95%.	65
Διάγραμμα 5.21: Τεστ του μοντέλου Πρόβλεψης της τιμής του ETH- μέθοδος LSTM.	66

Διάγραμμα 5.22: Μέρες εκπαίδευσης και δοκιμής του μοντέλου Random Forest για πρόβλεψη του ETH.	67
Διάγραμμα 5.23: Δοκιμή προβλέψεων τιμής ETH με το μοντέλο Random Forest.	67
Διάγραμμα 5.24: Αναπαράσταση ιστορικών στοιχείων χρονοσειράς για την τιμή του XRP. .	68
Διάγραμμα 5.25: Οπτικοποίηση δεδομένων χρονοσειράς μετά από διαφοροποίηση για επίτευξη σταθερότητας.	69
Διάγραμμα 5.26: Έλεγχος τυχαιότητας δεδομένων σε διάγραμμα ACF/ PACF.	69
Διάγραμμα 5.27: : Τεστ του μοντέλου Πρόβλεψης της τιμής του XRP- μέθοδος ARIMA.	70
Διάγραμμα 5.28: Πρόβλεψη της τιμής XRP για μία εβδομάδα (27/09/2022 - 4/10/2022) με διάστημα εμπιστοσύνης 95%.	70
Διάγραμμα 5.29: Τεστ του μοντέλου Πρόβλεψης της τιμής του XRP- μέθοδος LSTM.	72
Διάγραμμα 5.30: Μέρες εκπαίδευσης και δοκιμής του μοντέλου Random Forest για πρόβλεψη του XRP.	72
Διάγραμμα 5.31: Δοκιμή προβλέψεων τιμής XRP με το μοντέλο Random Forest.	73
Διάγραμμα 5.32: Αναπαράσταση ιστορικών στοιχείων χρονοσειράς για την τιμή του USDT.	74
Διάγραμμα 5.33: Έλεγχος τυχαιότητας δεδομένων σε διάγραμμα ACF.	75
Διάγραμμα 5.34: Έλεγχος τυχαιότητας δεδομένων σε διάγραμμα PACF.	75
Διάγραμμα 5.35: Τεστ του μοντέλου Πρόβλεψης της τιμής του USDT- μέθοδος ARIMA. ...	76
Διάγραμμα 5.36: Πρόβλεψη της τιμής USDT για μία εβδομάδα (27/09/2022 - 4/10/2022) με διάστημα εμπιστοσύνης 95%.	76
Διάγραμμα 5.37: Τεστ του μοντέλου Πρόβλεψης της τιμής του USDT- μέθοδος LSTM.	78
Διάγραμμα 5.38: Μέρες εκπαίδευσης και δοκιμής του μοντέλου Random Forest για πρόβλεψη του USDT.	78
Διάγραμμα 5.39: Δοκιμή προβλέψεων τιμής USDT με το μοντέλο Random Forest.	79
Διάγραμμα 5.40: Τιμή RMSE για πρόβλεψη του BTC με τρία διαφορετικά μοντέλα.	80
Διάγραμμα 5.41: Τιμή RMSE για πρόβλεψη του ETH με τρία διαφορετικά μοντέλα.	81
Διάγραμμα 5.42: Τιμή RMSE για πρόβλεψη του XRP με τρία διαφορετικά μοντέλα.	81
Διάγραμμα 5.43: Τιμή RMSE για πρόβλεψη του USDT με τρία διαφορετικά μοντέλα.	82

Κατάλογος Εξισώσεων

Εξίσωση 1: Εξίσωση υπολογισμού δείκτη MSE.	34
Εξίσωση 2: Εξίσωση υπολογισμού δείκτη RMSE.....	34
Εξίσωση 3: Εξίσωση υπολογισμού δείκτη AME.....	34
Εξίσωση 4: Εξίσωση υπολογισμού δείκτη MRE(%).	35
Εξίσωση 5: Εξίσωση πρόβλεψης για μοντέλο ARIMA(0,1,0).....	63

Κεφάλαιο 1

1 Εισαγωγή

Τα τελευταία χρόνια, η δραματική αύξηση του Bitcoin και των κρυπτονομισμάτων γενικά έχει προσελκύσει μεγάλη προσοχή. Αυξάνοντας περισσότερο από 120% το έτος 2016 και φτάνοντας στο «απίστευτο» επίπεδο των \$20.000 από \$900 το έτος 2017, οι τιμές του Bitcoin γνώρισαν μια εκθετική ανάπτυξη που οδήγησε σε ευκαιρίες για τεράστια κέρδη που καμία άλλη κατηγορία χρηματοοικονομικών περιουσιακών στοιχείων δεν κατάφερε σε τόσο σύντομο χρονικό διάστημα. Άλλα κρυπτονομίσματα όπως το Ethereum, το Ripple και το Litecoin δεν αποτέλεσαν εξαίρεση και οι τιμές τους αυξήθηκαν αρκετές χιλιάδες τοις εκατό μόνο το 2017 (Kriptomat, 2022).

Επιπλέον, η κυριαρχία του Bitcoin στην κεφαλαιοποίηση της αγοράς έναντι της αγοράς κρυπτονομισμάτων έχει σταδιακά εξασθενήσει από 85% το 2017 σε 50% σήμερα, δείχνοντας ότι μια συνολική έλξη για τα κρυπτονομίσματα έχει σημειωθεί τα τελευταία πέντε χρόνια.

Τον τελευταίο καιρό, καθώς το Bitcoin φαίνεται να ακολουθεί μία καθοδική πορεία, παρασύρει ολόκληρη την αγορά κρυπτονομισμάτων μαζί του. Οι συμμετέχοντες στην αγορά ενδιαφέρονται όλο και περισσότερο για τους παράγοντες που οδηγούν σε τέτοιες πτώσεις για να κατανοήσουν τη δυναμική των τιμών αυτών των ψηφιακών κρυπτονομισμάτων.

Ωστόσο, από την οπτική γωνία ενός εμπόρου κρυπτονομισμάτων, το αν οι τιμές ανεβαίνουν ή πέφτουν δεν είναι πρόβλημα, εφόσον η κατεύθυνση θεωρείται προβλέψιμη. Στην περίπτωση μιας αναμενόμενης περιόδου άνθησης, οι επενδυτές μπορούν να πάρουν θέση «long» στα κρυπτονομίσματα εκ των προτέρων για να αυξήσουν τις αποδόσεις τους μόλις οι τιμές φτάσουν σε ένα συγκεκριμένο επίπεδο. Ενώ στην περίπτωση μιας περιόδου ύφεσης που προβλέπεται στο μέλλον, οι επενδυτές μπορούν να πουλήσουν ανοικτά αυτά τα κρυπτονομίσματα μέσω διαπραγμάτευσης περιθωρίου (που επιτρέπεται από πολλά ανταλλακτήρια κρυπτονομισμάτων) για να κερδίσουν υπερβολικές αποδόσεις. Επιπλέον, η λήψη θέσεων «long» ή «short» έχει γίνει πολύ πιο εύκολη μετά τη δράση που ανέλαβε η CBOE τον Δεκέμβριο του 2017 όταν εισήγαγε τα συμβόλαια μελλοντικής εκπλήρωσης Bitcoin. Ένα τέτοιο χρηματοοικονομικό περιουσιακό

στοιχείο παρέχει στους επενδυτές να κάνουν εικασίες για τις τιμές Bitcoin και προς τις δύο κατευθύνσεις μέσω μόχλευσης χωρίς να κατέχουν Bitcoins (Akyildirim et al., 2020).

Η μεγάλη ανάπτυξη και οι συχνές διακυμάνσεις στις τιμές οδηγούν στο ερώτημα εάν οι τιμές των κρυπτονομισμάτων είναι προβλέψιμες. Με άλλα λόγια, ισχύει το Efficient Market Hypothesis (EMH) για τα κρυπτονομίσματα;

Κεφάλαιο 2

2 Ανασκόπηση βιβλιογραφίας

Το κρυπτονομίσματα είναι ένα νέο ψηφιακό περιουσιακό στοιχείο στα χρηματοοικονομικά, το οποίο έχει εξαιρετικά υψηλή μεταβλητότητα σε σύγκριση με σχεδόν όλα τα άλλα χρηματοοικονομικά μέσα. Ως αποτέλεσμα της υψηλής μεταβλητότητάς του και των διακυμάνσεων των τιμών, ένας πολύ περιορισμένος αριθμός άρθρων υπάρχει στη βιβλιογραφία που ασχολείται με την πρόβλεψη των τιμών (Chowdhury et al., 2020).

Οι κκ. (Lahmiri & Bekiros, 2019) έχουν προτείνει δύο τεχνικές βαθιάς μάθησης, συγκεκριμένα το νευρωνικό δίκτυο βαθιάς μάθησης (DLNN) και νευρωνικά δίκτυα γενικευμένης παλινδρόμησης (GRNN) για την πρόβλεψη της τιμής του Bitcoin, του Digital Cash και του Ripple. Στην περίπτωση του Bitcoin, το RMSE που λαμβάνεται με χρήση DLNN και GRNN είναι $2,75 \times 10^3$ και $8,80 \times 10^3$, αντίστοιχα. Στην περίπτωση του Digital Cash, αυτές οι τιμές είναι 19,2923 και 50,2418 και στην περίπτωση Ripple, 0,0499 και 0,3115. Φαίνεται ότι, στην περίπτωση του Bitcoin και του Digital Cash, η τιμή RMSE που λαμβάνεται με χρήση DLNN και GRNN είναι εξαιρετικά υψηλή, ενώ στην περίπτωση του Ripple, αυτή η τιμή είναι συγκριτικά χαμηλότερη από τις άλλες.

Οι κκ. (Greaves & Au, 2015) έχουν προτείνει δεδομένα γραφημάτων συναλλαγών για να προβλέψουν τις τιμές Bitcoin. Έχουν χρησιμοποιήσει τέσσερα μοντέλα ταξινόμησης: βασική γραμμή, λογιστική παλινδρόμηση, SVM και μοντέλο νευρωνικών δικτύων. Η ακρίβεια που επιτυγχάνεται χρησιμοποιώντας αυτά τα μοντέλα είναι 53,4% για τη γραμμή βάσης, 54,3% για την λογιστική παλινδρόμηση, 53,7% για το SVM και 55,1% για το νευρωνικό δίκτυο αντίστοιχα.

Οι κκ. (McNally et al., 2018) έχουν χρησιμοποιήσει επαναλαμβανόμενο νευρωνικό δίκτυο (RNN), δίκτυο μακράς βραχείας μνήμης (LSTM) και μοντέλο ARIMA για να προβλέψουν την κατεύθυνση της τιμής του Bitcoin σε USD. Το RNN και το LSTM είναι δύο αγωγοί βαθιάς μάθησης, ξεπέρασαν τις επιδόσεις του μοντέλου πρόβλεψης ARIMA. Το ριζικό μέσο τετράγωνο σφάλμα (RMSE) χρησιμοποιήθηκε για την αξιολόγηση και τη σύγκριση της ακρίβειας παλινδρόμησης και μια στρατηγική επικύρωσης 80/20 για την επικύρωση των μοντέλων. Ως αποτέλεσμα, η ακρίβεια και το RMSE που αποκτήθηκαν με το μοντέλο ARIMA

είναι 50,05% και 53,74%. Χρησιμοποιώντας το μοντέλο RNN (LSTM), η ακρίβεια και το RMSE που προέκυψαν είναι 50,25% (52,78%) και 5,45% (6,87%).

Υπάρχει ένας ικανοποιητικός αριθμός μελετών σχετικά με την αποτελεσματικότητα της τιμολόγησης του Bitcoin, ωστόσο άλλα κρυπτονομίσματα αγνοούνται ως επί το πλείστον σε αυτό το σκέλος της βιβλιογραφίας. Στην παρούσα εργασία θα γίνει προσπάθεια να καλυφθεί αυτό το ερευνητικό κενό. Τα μοντέλα πρόβλεψης και οι μεθοδολογίες εφαρμόζονται τόσο στο Bitcoin όσο και στα Ethereum, Ripple και Tether.

Κεφάλαιο 3

3 Εισαγωγή στα κρυπτονομίσματα

3.1 Εισαγωγή

Το κρυπτονομίσμα συναντάται στην αγορά με πολλά ονόματα. Πιθανότατα η πλειοψηφία των ανθρώπων έχει διαβάσει για μερικούς από τους πιο δημοφιλείς τύπους κρυπτονομισμάτων όπως το Bitcoin, το Litecoin και το Ethereum.

Σε αυτό το κεφάλαιο γίνεται αναφορά στα κρυπτονομίσματα και blockchain μέσα από επεξήγηση των ορισμών και της ιστορικής τους αναδρομής. Επιπλέον, σχολιάζονται τα δημοφιλέστερα παραδείγματα. Τέλος, για να κατανοηθεί περαιτέρω η λειτουργία τους τονίζονται τα μειονεκτήματα και πλεονεκτήματα της χρήσης τους καθώς και οι διαφορές τους από το ηλεκτρονικό χρήμα.

3.2 Κρυπτονομίσματα και blockchain

Τα τελευταία χρόνια έννοιες όπως τα κρυπτονομίσματα και τα blockchain έχουν εισέλθει στην ζωή μας. Πολλές φορές η διατύπωση των ορισμών είναι δύσκολη. Τα κρυπτονομίσματα είναι ψηφιακά ή αλλιώς εικονικά νομίσματα που χρησιμοποιούν ένα ηλεκτρονικό καθολικό (ένα είδος βιβλίου ή καταλόγου) με ισχυρή κρυπτογράφηση και από εκεί πηγάζει και το όνομά τους. Για τη διαδικασία της κρυπτογράφησης –δηλαδή της κωδικοποίησης των πληροφοριών– τα κρυπτονομίσματα χρησιμοποιούν το blockchain ως την υποκείμενη τεχνολογία για την καταγραφή, την επαλήθευση και την ασφάλεια των δεδομένων και των συναλλαγών.

Το πιο διάσημο και δημοφιλές κρυπτονομίσμα είναι αναμφίβολα το bitcoin, που κατάφερε να φτάσει, από την αξία του 1 δολαρίου το 2011 (ή ακόμα και από 0 δολάριο, όταν πρωτοεμφανίστηκε το 2009), στην κορυφή των σχεδόν 69.000 δολαρίων τον Νοέμβριο του 2021, ενώ πριν από λίγες ημέρες διαπραγματευόταν σε τιμή κατώτερη των 20.000 δολαρίων (Κριάρης, 2022).

Η τεχνολογία του blockchain με τη σειρά της πρόκειται για μια ψηφιακή βάση δεδομένων, η οποία αποτελείται από εγγραφές τέτοιων δεδομένων (ομαδοποιημένα ονομάζονται blocks) που συνδέονται μεταξύ τους κρυπτογραφικά σε μορφή αλυσίδας. Το Blockchain με άλλα λόγια αποτελεί ένα συγκεκριμένο τύπο ή υποσύνολο της λεγόμενης τεχνολογίας καταμεμημένου

καθολικού (« Distributed Ledger Technology (DLT)»). Το DLT είναι ένας τρόπος εγγραφής και κοινής χρήσης δεδομένων σε πολλαπλές αποθήκες δεδομένων (γνωστές και ως λογιστικά βιβλία), οι οποίες έχουν τα ίδια ακριβώς αρχεία δεδομένων και διατηρούνται και ελέγχονται συλλογικά από ένα καταναμημένο δίκτυο διακομιστών υπολογιστών, οι οποίοι ονομάζονται κόμβοι (Natarajan et al., 2017).

Το Blockchain χρησιμοποιεί μια μέθοδο κρυπτογράφησης με συγκεκριμένους μαθηματικούς αλγορίθμους για τη δημιουργία και την επαλήθευση μιας συνεχώς αναπτυσσόμενης δομής δεδομένων, όπου μπορούν μόνο να προστεθούν δεδομένα. Στην πράξη, το blockchain είναι μια τεχνολογία με πολλά «πρόσωπα». Μπορεί να εμφανίζει διαφορετικά χαρακτηριστικά και καλύπτει μια ευρεία γκάμα συστημάτων που κυμαίνονται από πλήρως ανοιχτά και χωρίς άδεια έως αδειοδοτημένα (Witzig & Salomon, 2019) :

- **Ανοιχτό blockchain (χωρίς άδεια):** Το άτομο μπορεί να εγγραφεί ή να αποχωρήσει από το δίκτυο κατά βούληση, χωρίς να χρειάζεται (προ)έγκριση από οποιαδήποτε (κεντρική) οντότητα. Για την ένταξη στο δίκτυο και την προσθήκη συναλλαγών στο καθολικό χρειάζεται μόνο ένας υπολογιστής στον οποίο έχει εγκατασταθεί το σχετικό λογισμικό. Δεν υπάρχει κεντρικός κάτοχος του δικτύου και του λογισμικού και πανομοιότυπα αντίγραφα του καθολικού διανέμονται σε όλους τους κόμβους του δικτύου. Η συντριπτική πλειονότητα των κρυπτονομισμάτων που κυκλοφορούν βασίζεται σε blockchains χωρίς άδεια (π.χ. Bitcoin, Litecoin,...) (Delahaye, 2016).
- **Αδειοδοτημένο blockchain:** Οι κόμβοι (συναλλάσσοντες) πρέπει να προεπιλεγούν από έναν διαχειριστή δικτύου, ο οποίος ορίζει τους κανόνες για το καθολικό. Έτσι, επιτρέπεται η εύκολη επαλήθευση της ταυτότητας των συμμετεχόντων του δικτύου. Ωστόσο, απαιτεί επίσης από τους συμμετέχοντες του δικτύου να εμπιστεύονται μια κεντρική οντότητα συντονισμού για την επιλογή αξιόπιστων κόμβων δικτύου. Γενικά, οι επιτρεπόμενες αλυσίδες μπλοκ μπορούν να χωριστούν περαιτέρω σε δύο υποκατηγορίες. Από τη μία πλευρά, υπάρχουν ανοιχτές ή δημόσιες εξουσιοδοτημένες αλυσίδες μπλοκ, στις οποίες μπορεί να έχει πρόσβαση και να τις δει οποιοσδήποτε, αλλά μόνο εξουσιοδοτημένοι συμμετέχοντες στο δίκτυο μπορούν να δημιουργήσουν συναλλαγές ή/και να ενημερώσουν την κατάσταση του καθολικού. Από την άλλη πλευρά, υπάρχουν κλειστές αλυσίδες μπλοκ ή blockchain με άδεια "επιχειρήσεων",

όπου η πρόσβαση είναι περιορισμένη και όπου μόνο ο διαχειριστής του δικτύου μπορεί να δημιουργήσει συναλλαγές και να ενημερώσει την κατάσταση του καθολικού. Αυτό που είναι σημαντικό να σημειωθεί είναι ότι ακριβώς όπως σε ένα ανοιχτό blockchain χωρίς άδεια, οι συναλλαγές σε ένα blockchain με ανοιχτή άδεια μπορούν να επικυρωθούν και να εκτελεστούν χωρίς τη μεσολάβηση ενός αξιόπιστου τρίτου μέρους. Ορισμένα κρυπτονομίσματα, όπως το Ripple και το NEO, χρησιμοποιούν blockchains με δημόσια άδεια (Delahaye, 2016).

3.3 Ιστορική αναδρομή

Η ιδέα πίσω από τα κρυπτονομίσματα

Το κρυπτονόμισμα εμφανίστηκε για πρώτη φορά στις αρχές της δεκαετίας του 1980, όταν ο Αμερικανός κρυπτογράφος David Chaum δημοσίευσε μια εργασία συνεδρίου που περιγράφει μια πρώιμη μορφή ανώνυμου κρυπτογραφικού ηλεκτρονικού χρήματος. Η ιδέα πίσω από αυτό το νόμισμα είναι ότι μπορεί να σταλεί χωρίς παρακολούθηση και χωρίς χρήση κεντρικών οντοτήτων. Το 1995, ο Chaum ανέπτυξε ένα πρωτο-κρυπτονόμισμα που ονομάζεται Digicash. Χρησιμοποίησε κρυπτογραφικούς αλγόριθμους για να δημιουργήσει ένα ασφαλές ψηφιακό σύστημα για την ανταλλαγή χρημάτων. Ήταν απαραίτητο για τον χρήστη να εγκαταστήσει λογισμικό που θα μπορούσε να κάνει ανάληψη χρημάτων από μια τράπεζα και απαιτούσε συγκεκριμένα κρυπτογραφημένα κλειδιά για να μπορέσουν να σταλούν τα χρήματα στον προβλεπόμενο παραλήπτη.

Ο βιβλιογραφικός πρόδρομος του BTC, το Bit Gold δημιουργήθηκε το 1998 από τον Nick Szabo (Burke, 2021). Ήταν απαραίτητο για έναν συμμετέχοντα να αφιερώσει την ισχύ του υπολογιστή για να λύσει κρυπτογραφικούς γρίφους και όσοι έλυναν τους γρίφους έπαιρναν ανταμοιβή. Ο συνδυασμός της δουλειάς του Chaum με αυτό οδηγεί σε κάτι που μοιάζει πολύ με το Bitcoin.

Αλλά ο Szabo δεν θα μπορούσε να λύσει το περιβόητο πρόβλημα της διπλής δαπάνης (τα ψηφιακά δεδομένα μπορούν να αντιγραφούν και να επικολληθούν) χωρίς τη χρήση μιας κεντρικής αρχής. Μια δεκαετία αργότερα, ένα μυστηριώδες άτομο ή ομάδα, χρησιμοποιώντας το ψευδώνυμο Satoshi Nakamoto, δημοσίευσε μια λευκή βίβλο με τίτλο «Bitcoin – A Peer to Peer Electronic Cash System» (Hayes, 2022). Αυτό το έγγραφο έθεσε σε κίνηση την ιστορία του Bitcoin και αργότερα των κρυπτονομισμάτων.

Αρχή (2008-2010)

Στις 31 Οκτωβρίου 2008, ο Satoshi Nakamoto δημοσίευσε τη λευκή βίβλο που περιγράφει τη λειτουργικότητα του δικτύου blockchain Bitcoin (Burke, 2021). Ο Satoshi Nakamoto ξεκίνησε επίσημα να εργάζεται στο έργο bitcoin στις 18 Αυγούστου 2008, όταν αγόρασε το Bitcoin.org. Το Bitcoin και άλλα κρυπτονομίσματα δεν θα ήταν δυνατά χωρίς την τεχνολογία blockchain, η οποία είναι ένας απλός τρόπος δημιουργίας αμετάβλητων δομών δεδομένων.

Η ιστορία του Bitcoin βρισκόταν σε εξέλιξη. Ο Satoshi Nakamoto εξόρυξε το πρώτο μπλοκ του δικτύου Bitcoin στις 3 Ιανουαρίου 2009. Ενσωμάτωσαν έναν τίτλο από την εφημερίδα The Times σε αυτό το αρχικό μπλοκ, κάνοντας μια μόνιμη αναφορά στις οικονομικές συνθήκες - που περιλαμβάνουν προγράμματα διάσωσης τραπεζών και ένα κεντρικό χρηματοπιστωτικό σύστημα - ότι το Bitcoin ήταν εν μέρει μια αντίδραση κατά.

Το πρώτο μπλοκ που είχε ως αποτέλεσμα την εξόρυξη 50 bitcoin είναι τώρα γνωστό ως Genesis Block (Burke, 2021). Το Bitcoin είχε πολύ μικρή αξία αυτή τη στιγμή, όπως είχε τους πρώτους μήνες της ύπαρξής του. Έξι μήνες αφότου το bitcoin έγινε εμπορεύσιμο, τον Απρίλιο του 2010, ένα bitcoin άξιζε περίπου 0,14 \$. Στις αρχές Νοεμβρίου, η τιμή του bitcoin εκτινάχθηκε στα 36 σεντς πριν διαμορφωθεί γύρω στα 29 σεντς.

Σχηματισμός της αγοράς (2010-2014)

Παρόλο που η αξία του ήταν ακόμα πολύ μικρή, το Bitcoin έδειχνε ότι είχε πραγματική παγκόσμια αξία. Τον Φεβρουάριο του 2011 ανέβηκε στα 1,06 δολάρια πριν υποχωρήσει στα 87 σεντς περίπου. Την άνοιξη, εν μέρει λόγω μιας ιστορίας του Forbes για το νέο «κρυπτονομίσμα», η τιμή απογειώθηκε. Από τις αρχές Απριλίου έως τα τέλη Μαΐου, η αξία του bitcoin αυξήθηκε από 86 σεντς στα 8,89 δολάρια.

Την 1η Ιουνίου, αφότου ο Gawker δημοσίευσε μια ιστορία σχετικά με την απήχηση του νομίσματος στην διαδικτυακή κοινότητα διακίνησης ναρκωτικών, η τιμή υπερτριπλασιάστηκε σε μια εβδομάδα, στα 27 \$ περίπου (Schurman, 2012). Η αγοραία αξία των bitcoins σε κυκλοφορία πλησίασε σχεδόν τα 130 εκατομμύρια δολάρια. Ωστόσο, μέχρι τον Σεπτέμβριο του 2011, η αξία είχε πέσει ξανά στα 4,77 \$ περίπου.

Τον Οκτώβριο του ίδιου έτους εμφανίστηκε το Litecoin (CoinLoan, 2020), ένα από τα πολλά forks (δηλαδή ενημερωμένες εκδόσεις) του Bitcoin. Σύντομα το Litecoin έγινε το δεύτερο μεγαλύτερο κρυπτονομίσμα ανά κεφαλαιοποίηση αγοράς, με το αρχαιότερο αρχείο του

CoinMarketCap (από τον Μάιο του 2013) που δείχνει τα PPCoin, Namecoin και 10 άλλα να ακολουθούν σε απόσταση. Τέτοια κρυπτονομίσματα ονομάστηκαν γρήγορα «altcoins», με μερικά να διαχωρίζονται από το Bitcoin και άλλα να βασίζονται σε νέο κώδικα.

Το 2012, οι τιμές του Bitcoin αυξάνονταν σταθερά και τον Σεπτέμβριο του ίδιου έτους ιδρύθηκε το Ίδρυμα Bitcoin για να προωθήσει την ανάπτυξη και την απορρόφηση του Bitcoin. Τότε γνωστό ως OpenCoin, το Ripple κυκλοφόρησε επίσης εκείνο το έτος, με το έργο να προσελκύει επιχειρηματικά κεφάλαια τον επόμενο χρόνο .

Το 2013, εν μέσω ομοσπονδιακών, ποινικών, κανονιστικών θεμάτων και ζητημάτων που σχετίζονται με το λογισμικό, η τιμή του bitcoin συνεχώς αυξανόταν και έπεσε ραγδαία. Στις 19 Νοεμβρίου η τιμή του έφτασε τα 755 δολάρια, για να πέσει στα 378 δολάρια την ίδια μέρα. Μέχρι τις 30 Νοεμβρίου ήταν και πάλι στα 1.163 \$. Αυτή ήταν, ωστόσο, η αρχή ενός άλλου μακροπρόθεσμου κραχ που έληξε με το Bitcoin να πέφτει ξανά στα 152 \$ έως τον Ιανουάριο του 2015.

Οι απάτες (2014-2016)

Αν και σκόπιμη, η ανωνυμία και η έλλειψη κεντρικού ελέγχου καθιστούν το ψηφιακό νόμισμα ιδιαίτερα ελκυστικό για τους εγκληματίες. Το Φεβρουάριο του 2014, το Mt.Gox — τότε το μεγαλύτερο ανταλλακτήριο bitcoin στον κόσμο, με έδρα την Ιαπωνία — κατέρρευσε και κήρυξε πτώχευση, έχοντας χάσει 850.000 bitcoin. Αν και δεν είναι γνωστό τι ακριβώς συνέβη, είναι πιθανό ότι το BTC που έλειπαν κλάπηκαν σταδιακά με την πάροδο του χρόνου, ξεκινώντας το 2011, και μεταπωλήθηκαν σε διάφορα ανταλλακτήρια για μετρητά (συμπεριλαμβανομένου του Mt.Gox) (Tuwiner, 2022).

Αν και μικρότερα χρηματιστήρια συνεχίζουν να παραβιάζονται ακόμη και σήμερα, οι μεγαλύτερες πλατφόρμες παρέχουν πλέον περισσότερες εγγυήσεις για τα αποθεματικά τους σε περίπτωση παραβιάσεων. Αυτό περιλαμβάνει το Secure Asset Fund for Users on Binance, για παράδειγμα, το οποίο είναι ένα ασφαλιστικό ταμείο έκτακτης ανάγκης. Συνιστάται στους εμπόρους κρυπτονομισμάτων να χρησιμοποιούν ένα πορτοφόλι υλικού ή λογισμικού για την ασφαλή αποθήκευση των κρυπτονομισμάτων τους αντί να τα αποθηκεύουν σε ένα χρηματιστήριο. Πορτοφόλια όπως αυτά δεν ήταν τόσο προσβάσιμα κατά τη διάρκεια αυτής της πρώιμης περιόδου στην ιστορία των κρυπτονομισμάτων.

Το Bitcoin ένα παγκόσμιο φαινόμενο (2016-2018)

Οι τιμές του Bitcoin αυξήθηκαν σταθερά σε ετήσια βάση, από 434 \$ τον Ιανουάριο του 2016 σε 998 \$ τον Ιανουάριο του 2017. Τον Ιούλιο του 2017, εγκρίθηκε μια αναβάθμιση λογισμικού σε Bitcoin, με στόχο να υποστηρίξει την ανάπτυξη του Δικτύου Lightning καθώς και βελτίωση της ασφάλειας .

Μια εβδομάδα μετά την ενεργοποίηση της αναβάθμισης τον Αύγουστο, το Bitcoin διαπραγματευόταν περίπου στα 2.700 δολάρια. Μέχρι τις 17 Δεκεμβρίου 2017, το Bitcoin έφτασε στο αστρονομικό υψηλό όλων των εποχών λίγο κάτω από τα \$20.000 (Jones, 2021).

Την ίδια περίοδο, ένα νέο έργο blockchain που ονομάζεται Ethereum έκανε θόρυβο στη σφαίρα των κρυπτονομισμάτων, έχοντας γίνει γρήγορα το νούμερο δύο στα κρυπτονομίσματα ανά κεφαλαιοποίηση αγοράς από τότε που κυκλοφόρησε τον Ιούλιο του 2015. Έφερε έξυπνα συμβόλαια στα κρυπτονομίσματα, ανοίγοντας ένα ευρύ φάσμα πιθανών περιπτώσεων χρήσης και δημιουργώντας πάνω από 200.000 διαφορετικά πρότζεκτ (και συνεχίζονται). Σε αντίθεση με το Bitcoin, το Ethereum επιτρέπει σε πρόσθετες πλατφόρμες να ξεκινούν και να λειτουργούν στη δική του αλυσίδα, η καθεμία με τα δικά της κρυπτονομίσματα και τις δικές της περιπτώσεις χρήσης. Αυτό ήταν ένα μοντέλο που αντιγράφηκε ευρέως από άλλα νέα blockchains, με Cardano, Tezos και Neo που κυκλοφόρησαν κατά τη διάρκεια αυτής της περιόδου (Marr, 2021).

Αυξομειώσεις (2018 έως Σήμερα)

Το Bitcoin δεν μπόρεσε να διατηρήσει την υψηλή τιμή όλων των εποχών των 19.783 \$. Ομοίως, το Ethereum, το οποίο έφτασε στο δικό του πικ τον Ιανουάριο του 2018 περίπου στα 1.400 \$, δεν μπόρεσε επίσης να διατηρήσει το νεοανακαλυφθέν επίπεδό του για πολύ. Οι οικονομικοί κανονισμοί και οι ανησυχίες για την ασφάλεια (λόγω ημι-κανονικών εισβολών συναλλάγματος) συνέβαλαν στην πτώση σε όλη την αγορά και μέχρι το τέλος του 2018 το bitcoin είχε πέσει σε περίπου 3.700 δολάρια.

Ωστόσο, οι τιμές δεν παρέμειναν χαμηλές για πολύ καιρό. Ξεκινώντας από τα τέλη του 2020, το bitcoin γνώρισε κάτι σαν αναγέννηση, ξεκινώντας με την ανακοίνωση της MicroStrategy («εταιρεία επιχειρηματικής ευφυΐας») τον Αύγουστο ότι είχε αγοράσει bitcoin αξίας 250 εκατομμυρίων δολαρίων . Αυτό ξεκίνησε μια ανοδική αγορά στην οποία προστέθηκαν και οι υπόλοιπες αγορές, με τις τιμές να ενισχύθηκαν περαιτέρω από την αγορά 1,5 δισεκατομμυρίων

δολαρίων σε bitcoin από την Tesla στις αρχές του 2021. Το Νοέμβριο του ίδιου έτους, το bitcoin έφτασε στο τρέχον υψηλό ρεκόρ των 69.000 δολαρίων (Jones, 2021).

Η αγορά έπεσε για άλλη μια φορά μετά από αυτό το υψηλό, παρασυρμένη από τις μακροοικονομικές ανησυχίες που προκύπτουν από τον υψηλό πληθωρισμό, την αύξηση των επιτοκίων και το φάσμα του πολέμου. Με τις παγκόσμιες χρηματιστηριακές αγορές να πέφτουν επίσης στα τέλη του 2021 και το 2022, η παράλληλη πτώση της κρυπτογράφησης δείχνει ότι ο κλάδος περιπλέκεται όλο και περισσότερο με τις παραδοσιακές χρηματοπιστωτικές αγορές.

Και ενώ η αστάθεια των κρυπτονομισμάτων είναι τόσο ελκυστική όσο και δυνητικά καταστροφική, η υποκείμενη τεχνολογία πίσω από όλα αυτά, το blockchain, έχει τη δύναμη να αλλάξει πολλούς τομείς της κοινωνίας μας. Είτε παρέχει προσιτές επιλογές χρηματοοικονομικών συναλλαγών, διασφαλίζοντας τα ίδια κεφάλαια, είτε παρέχει ακριβή δεδομένα για την ασφαλιστική προσφορά, η τεχνολογία blockchain έχει τη δυνατότητα να χρησιμοποιηθεί σχεδόν σε κάθε τομέα της οικονομίας.

Καθώς η αγορά γίνεται πιο σταθερή με αυξημένη γνώση και με την εισαγωγή νέων τομέων όπως τα stablecoins και την αποκεντρωμένη χρηματοδότηση (DeFi), είναι εύκολο να ενθουσιαστεί κανείς με τα κρυπτονομίσματα και τις δυνατότητές τους από επενδυτική και τεχνολογική προοπτική.

3.4 Δημοφιλέστερα κρυπτονομίσματα και blockchain

Οι επενδυτές αναλογίζονται για κερδοφόρα νομίσματα, διακριτικά και μη για τα δίκτυα blockchain που τα υποστηρίζουν. Ο ενδοιασμός των επενδυτών είναι απολύτως θεμιτός αν αναλογιστεί κανείς ότι η κατανόηση της τεχνολογίας blockchain ήταν κρίσιμη στις πρώτες εμφανίσεις των κρυπτονομισμάτων. Σήμερα, είναι λογικό να γίνονται επενδύσεις χωρίς να συνυπολογίζεται το θέμα των χρησιμοποιημένων blockchain. Παρόλα αυτά, τα blockchain είναι η βάση κάθε σύγχρονου κρυπτονομίσματος όσο και ως προς τους τρόπους που διαφέρουν μεταξύ τους. Σε αυτή την ενότητα, ακολουθεί μια περιήγηση στα πιο ευρέως χρησιμοποιούμενα blockchain της βιομηχανίας των κρυπτονομισμάτων καθώς και στα πιο δημοφιλή νομίσματα της αγοράς.

Πολλές φορές, το blockchain με το αντίστοιχο κρυπτονομισμά του είναι στενά συνδεδεμένα και αναφέρονται σαν ένα. Παραδείγματα στενής σύνδεσης μεταξύ blockchain και νομισμάτων αποτελούν: το Bitcoin blockchain (το θεμέλιο του Bitcoin), Ethereum blockchain (Ether),

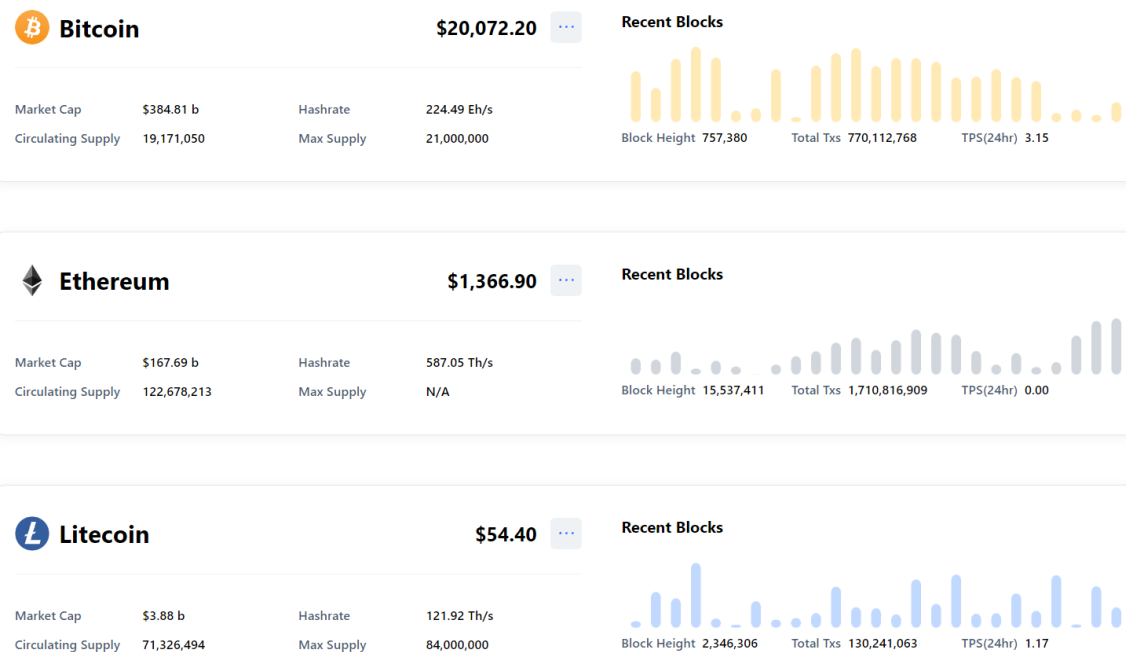
το Litecoin, το XRP, το Eos, το Tron, το Solana, το Stellar, το Neo, το Monero, το Dogecoin lockchain.

Ορισμένα κρυπτονομίσματα εφαρμόζονται σε υπάρχουσες αλυσίδες blockchain. Κλασσικό παράδειγμα αποτελεί το Ethereum blockchain όπου πολλά νομίσματα βασίζονται σε αυτό, όπως: το Chainlink, το Maker, το Status, το Uniswap, το Axie Infinity, το Aave, το Compound, το SushiSwap, το Kyber Network, το Basic Attention Token και το Decentraland.

Όλα τα κρυπτονομίσματα χρησιμοποιούν μεθόδους κρυπτογράφησης blockchain για να διασφαλίσουν ότι τα δεδομένα είναι ασφαλή και οι συναλλαγές επικυρώνονται. Ωστόσο, υπάρχει διαφορά μεταξύ των κρυπτονομισμάτων που έχουν τις δικές τους αλυσίδες και εκείνων που φιλοξενούνται σε υπάρχοντα δίκτυα blockchain. Όταν ένα κρυπτονομίσμα είναι το εγγενές νόμισμα του blockchain του, αναφερόμαστε στις μονάδες ονομαστικής αξίας του ως «νομίσματα» (coin). Τα Ether, NEO και Dogecoin είναι όλα νομίσματα. Όταν το κρυπτονομίσμα φιλοξενείται σε ένα υπάρχον blockchain, αναφερόμαστε πλέον σε ένα «διακριτικό» (token). Τα MakerDAO, Chainlink και BAT είναι όλα διακριτικά επειδή φιλοξενούνται στο Ethereum και δεν έχουν δικά τους blockchain.

3.4.1 Κορυφαία Δίκτυα Blockchain

Το Bitcoin (BTC) αντιπροσωπεύει περίπου το 40% της συνολικής αξίας των κρυπτονομισμάτων. Οπότε είναι λογική η υπόθεση ότι το Bitcoin blockchain είναι το πιο διαδεδομένο και αναμφισβήτητο αυτό που έχει συγκεντρώσει τη μεγαλύτερη αξία.













Εικόνα 3.1: Αξία των δημοφιλέστερων δικτύων blockchain της αγοράς, Πηγή (6/10.2022): https://blockchain.coinmarketcap.com/?utm_source=coinmarketcap&utm_content=footer.

Παρολ'αυτά δεν είναι λάθος ο ισχυρισμός ότι το Ethereum blockchain είναι εξίσου δημοφιλές. Για να κριθεί η δημοτικότητα τους πρέπει να υπολογιστούν και άλλοι παράμετροι πέρα από την αξία στην αγορά. Το 2021, ο συνολικός αριθμός των συναλλαγών στο Ethereum ξεπέρασε για πρώτη φορά τον αριθμό των συναλλαγών στο Bitcoin (Kriptomat, 2022). Επειδή το Ethereum είναι το δίκτυο θεμελίωσης για πολλά κρυπτονομίσματα (και τα περισσότερα NFTs) είναι λογικό να προσδιορίσουμε το Ethereum ως το πιο δημοφιλές δίκτυο blockchain στον κόσμο.

Παρά την φήμη των δύο προαναφερόμενων blockchain δεν πρέπει να υποτιμάται και η δύναμη των υπολοίπων. Για παράδειγμα, το Solana blockchain προορίζεται να χρησιμεύσει ως δίκτυο θεμελίωσης για κρυπτονομίσματα και κατανεμημένες εφαρμογές. Το βασικό πλεονέκτημα του Solana είναι ότι ο μοναδικός μηχανισμός συναίνεσης Proof of History επιτρέπει στο δίκτυο να επεξεργάζεται δεκάδες χιλιάδες συναλλαγές ανά δευτερόλεπτο. Αυτό είναι ένα σημαντικό ζήτημα για ένα δίκτυο παγκόσμιας εμβέλειας. Για συγκριτικούς λόγους αναφέρεται ότι το Bitcoin υποστηρίζει 7 tps (συναλλαγές ανά δευτερόλεπτο) και το Ethereum υποστηρίζει περίπου 15 tps. Επιπλέον, αυτή η καινοτόμος αλυσίδα χρεώνει πολύ λιγότερα τέλη συναλλαγής από το Ethereum. Δεν είναι περίεργο που αυξάνεται ραγδαία σε δημοτικότητα.

3.4.2 Κορυφαία Κρυπτονομίσματα

Στις μέρες μας υπάρχουν πολλά κρυπτονομίσματα που μπορούν να αξιοποιηθούν για επενδύσεις. Παρακάτω αναφέρονται τα δέκα κυριότερα λαμβάνοντας υπόψιν την σημερινή αγοραστική δύναμη και την τιμή τους (9/2022). Πηγή της σημερινής και παλιότερης τιμής του κεφαλαίου αγοράς για το κάθε κρυπτόνμισμα αποτελεί η ιστοσελίδα: <https://coinmarketcap.com/> και η ανάταξη πραγματοποιήθηκε στις 6/10/2022.

#	Name	Price	1h %	24h %	7d %	Market Cap	Volume(24h)	Circulating Supply	Last 7 Days
☆ 1	 Bitcoin BTC	\$20,049.87	▼0.24%	▼0.82%	▲3.18%	\$384,394,305,356	\$36,019,563,432 1,796,434 BTC	19,171,218 BTC	
☆ 2	 Ethereum ETH	\$1,359.94	▼0.32%	▲0.29%	▲1.93%	\$166,743,639,442	\$12,053,500,498 8,868,288 ETH	122,680,595 ETH	
☆ 3	 Tether USDT	\$1.00	▼0.01%	▼0.00%	▲0.00%	\$68,223,803,250	\$44,391,863,390 44,384,557,032 USDT	68,212,574,446 USDT	
☆ 4	 BNB BNB	\$293.21	▼0.34%	▼0.31%	▲3.73%	\$47,303,782,894	\$634,705,087 2,164,765 BNB	161,337,261 BNB	
☆ 5	 USD Coin USDC	\$1	▼0.01%	▼0.01%	▼0.02%	\$46,255,939,417	\$3,911,032,423 3,910,708,404 USDC	46,252,107,237 USDC	

Εικόνα 3.2: Τα 5 πιο κορυφαία κρυπτονομίσματα ανά κεφαλαιοποίηση αγοράς. Πηγή (6/10/2022): <https://coinmarketcap.com/>

1. Bitcoin (BTC) - Κεφάλαιο αγοράς: 384 δισεκατομμύρια δολάρια

Δημιουργήθηκε το 2009 από τον Satoshi Nakamoto, το Bitcoin (BTC) είναι το πρώτο κρυπτόνμισμα. Η τιμή του Bitcoin έχει εκτοξευθεί καθώς έχει γίνει γνωστό όνομα (Tretina & Powell, 2022). Τον Μάιο του 2016, θα μπορούσατε να αγοράσετε ένα Bitcoin για περίπου 500 \$. Από τις 30 Σεπτεμβρίου 2022, η τιμή ενός Bitcoin ήταν περίπου 19.531 \$. Πρόκειται για αύξηση άνω του 3.800% .

2. Ethereum (ETH)- Κεφάλαιο αγοράς: 167 δισεκατομμύρια δολάρια

Τόσο ένα κρυπτόνμισμα όσο και μια πλατφόρμα blockchain, το Ethereum είναι το αγαπημένο κρυπτόνμισμα των προγραμματιστών λόγω των πιθανών εφαρμογών του, όπως τα λεγόμενα έξυπνα συμβόλαια που εκτελούνται αυτόματα όταν πληρούνται οι προϋποθέσεις και τα μη ανταλλάξιμα token (NFTs). Το Ethereum γνώρισε επίσης τεράστια ανάπτυξη. Από τον

Απρίλιο του 2016 έως το τέλος Σεπτεμβρίου 2022, η τιμή του αυξήθηκε από περίπου 11 \$ σε περίπου 1.388 \$, αυξάνοντας περίπου 12.518%.

3. Tether (USDT) - Κεφάλαιο αγοράς: 67 δισεκατομμύρια δολάρια

Σε αντίθεση με ορισμένες άλλες μορφές κρυπτονομισμάτων, το Tether (USDT) είναι ένα stablecoin, που σημαίνει ότι υποστηρίζεται από νομίσματα fiat όπως τα δολάρια ΗΠΑ και το ευρώ και υποθετικά διατηρεί μια αξία ίση με μία από αυτές τις ονομαστικές αξίες. Θεωρητικά, αυτό σημαίνει ότι η αξία του Tether υποτίθεται ότι είναι πιο συνεπής από άλλα κρυπτονομίσματα και ευνοείται από επενδυτές που είναι επιφυλακτικοί για την ακραία αστάθεια άλλων νομισμάτων.

4. Νόμισμα Binance (BNB)- Κεφάλαιο αγοράς: 47 δισεκατομμύρια δολάρια

Το Binance Coin (BNB) είναι μια μορφή κρυπτονομίσματος που μπορείτε να χρησιμοποιήσετε για συναλλαγές και πληρωμή προμηθειών στο Binance, ένα από τα μεγαλύτερα ανταλλακτήρια κρυπτογράφησης στον κόσμο. Από την κυκλοφορία του το 2017, το Binance Coin έχει επεκταθεί στο παρελθόν απλώς διευκολύνοντας τις συναλλαγές στην πλατφόρμα ανταλλαγής του Binance. Τώρα, μπορεί να χρησιμοποιηθεί για συναλλαγές, διεκπεραίωση πληρωμών ή ακόμα και για κρατήσεις ταξιδιών. Μπορεί επίσης να ανταλλάσσεται ή να ανταλλάσσεται με άλλες μορφές κρυπτονομισμάτων, όπως το Ethereum ή το Bitcoin. Η τιμή της BNB το 2017 ήταν μόλις 0,10 \$. Μέχρι τα τέλη Σεπτεμβρίου 2022, η τιμή του είχε αυξηθεί σε περίπου \$283, ένα κέρδος περίπου 282.900%.

5. Νόμισμα Δολάριο ΗΠΑ (USDC)- Κεφάλαιο αγοράς: 47 δισεκατομμύρια δολάρια

Όπως το Tether, το USD Coin (USDC) είναι ένα stablecoin, που σημαίνει ότι υποστηρίζεται από δολάρια ΗΠΑ και στοχεύει σε αναλογία 1 USD προς 1 USDC. Το USDC τροφοδοτείται από το Ethereum και μπορείτε να χρησιμοποιήσετε το USD Coin για να ολοκληρώσετε παγκόσμιες συναλλαγές.

6. XRP (XRP) - Κεφάλαιο αγοράς: 24 δισεκατομμύρια δολάρια

Δημιουργημένο από ορισμένους από τους ίδιους ιδρυτές με την Ripple, μια εταιρεία ψηφιακής τεχνολογίας και επεξεργασίας πληρωμών, το XRP μπορεί να χρησιμοποιηθεί σε αυτό το δίκτυο για να διευκολύνει τις ανταλλαγές διαφορετικών τύπων νομισμάτων, συμπεριλαμβανομένων των νομισμάτων fiat και άλλων σημαντικών κρυπτονομισμάτων. Στις αρχές του 2017, η τιμή

του XRP ήταν 0,006 \$. Στις 30 Σεπτεμβρίου 2022, η τιμή του έφτασε τα 0,48 \$, ισοδυναμώντας με άνοδο 7.900%.

7. Binance USD (BUSD)- Κεφάλαιο αγοράς: 21 δισεκατομμύρια δολάρια

Το Binance USD (BUSD) είναι ένα stablecoin που ίδρυσαν οι Paxos και Binance με σκοπό την δημιουργία ενός κρυπτονομίσματος που υποστηρίζεται από το δολάριο ΗΠΑ. Για να διατηρήσει αυτή την αξία, οι Paxos κατέχουν ποσό δολαρίων ΗΠΑ ίσο με τη συνολική προσφορά BUSD. Όπως και με άλλα stablecoins, το BUSD δίνει στους εμπόρους και στους χρήστες κρυπτονομισμάτων τη δυνατότητα να συμμετέχουν σε συναλλαγές με άλλα περιουσιακά στοιχεία κρυπτογράφησης, ελαχιστοποιώντας τον κίνδυνο αστάθειας.

8. Cardano (ADA)- Κεφάλαιο αγοράς: 14 δισεκατομμύρια δολάρια

Λίγο αργότερα στη σκηνή της κρυπτογράφησης, το Cardano (ADA) είναι αξιοσημείωτο για την πρόωπη αγκάλιασμα της επικύρωσης απόδειξης στοιχείου. Αυτή η μέθοδος επιταχύνει τον χρόνο συναλλαγής και μειώνει τη χρήση ενέργειας και τις περιβαλλοντικές επιπτώσεις καταργώντας την ανταγωνιστική πτυχή επίλυσης προβλημάτων της επαλήθευσης συναλλαγών σε πλατφόρμες όπως το Bitcoin. Το Cardano λειτουργεί επίσης όπως το Ethereum για να ενεργοποιήσει τα έξυπνα συμβόλαια και τις αποκεντρωμένες εφαρμογές, τις οποίες εξουσιοδοτεί η ADA, το εγγενές της νόμισμα.

Το διακριτικό ADA του Cardano είχε σχετικά μέτρια ανάπτυξη σε σύγκριση με άλλα μεγάλα κρυπτονομίσματα. Το 2017, η τιμή της ADA ήταν 0,02 \$. Στις 30 Σεπτεμβρίου 2022, η τιμή του ήταν 0,44 \$. Πρόκειται για αύξηση 2.100%.

9. Σολάνα (SOL)- Κεφάλαιο αγοράς: 12 δισεκατομμύρια δολάρια

Αναπτύχθηκε για να βοηθήσει τις χρήσεις αποκεντρωμένης χρηματοδότησης (DeFi), τις αποκεντρωμένες εφαρμογές (DApps) και τα έξυπνα συμβόλαια. Το Solana τρέχει με έναν μοναδικό υβριδικό μηχανισμό απόδειξης συμμετοχής και απόδειξης του ιστορικού για την επεξεργασία συναλλαγών γρήγορα και με ασφάλεια. Το εγγενές διακριτικό του, τροφοδοτεί την πλατφόρμα. Όταν κυκλοφόρησε το 2020, η τιμή του SOL ξεκίνησε από 0,77 \$. Μέχρι τα τέλη Σεπτεμβρίου 2022, η τιμή του ήταν περίπου 33,72 \$, ένα κέρδος σχεδόν 4.279%.

10. Dogecoin (DOGE)- Κεφάλαιο αγοράς: 8 δισεκατομμύρια δολάρια

Το Dogecoin ξεκίνησε ως πείραμα το 2013, αλλά εξελίχθηκε γρήγορα σε ένα εξέχον κρυπτονομίσμα χάρη σε μια αφοσιωμένη κοινότητα και δημιουργικά μιμίδια. Σε αντίθεση με πολλά άλλα κρυπτονομίσματα, δεν υπάρχει όριο στον αριθμό των Dogecoins που μπορούν να δημιουργηθούν, γεγονός που αφήνει το νόμισμα επιρρεπές σε υποτίμηση καθώς αυξάνεται η προσφορά. Η τιμή του Dogecoin το 2017 ήταν 0,0002 \$. Μέχρι τις 30 Σεπτεμβρίου 2022, η τιμή του ήταν στα 0,06 \$, αυξημένη κατά 29.900%.

3.5 Πλεονεκτήματα και μειονεκτήματα των κρυπτονομισμάτων

Τα κρυπτονομίσματα έχουν τη δυνατότητα να αναδιαμορφώσουν τον ήδη γνωστό οικονομικό κόσμο και να αμφισβητούν την ίδια την ύπαρξη της παραδοσιακής χρηματοοικονομικής υποδομής. Σύγχρονα ερωτήματα δημιουργούνται στους επενδυτές σχετικά με το αν αξίζει ή όχι η επένδυση σε κρυπτονομίσματα και αν ναι σε ποια. Για να απαντήσει κανείς στα παραπάνω ερωτήματα πρέπει να αναλογιστεί κανείς τα πλεονεκτήματα και τα μειονεκτήματα μιας τέτοιας επένδυσης. Σε αυτή την υποενότητα αναφέρονται μερικά από τα κυριότερα.

Πλεονεκτήματα

- Τα κρυπτονομίσματα αντιπροσωπεύουν ένα νέο, αποκεντρωμένο είδος χρήματος. Σε αυτό το σύστημα λείπουν οι κεντρικοί μεσάζοντες, όπως οι τράπεζες και τα νομισματικά ιδρύματα. Έτσι, ένα σύστημα με κρυπτονομίσματα εξαλείφει την πιθανότητα ενός μόνο σημείου πτώχευσης, όπως μια μεγάλη τράπεζα, πυροδοτώντας έναν καταρράκτη κρίσεων σε όλο τον κόσμο, όπως αυτή που προκλήθηκε το 2008 από την αποτυχία ιδρυμάτων στις Ηνωμένες Πολιτείες (Frankenfield, 2022).
- Τα κρυπτονομίσματα καλούνται να διευκολύνουν την απευθείας μεταφορά κεφαλαίων μεταξύ δύο μερών, χωρίς την ανάγκη ενός αξιόπιστου τρίτου, όπως μια τράπεζα ή μια εταιρεία πιστωτικών καρτών. Τέτοιες αποκεντρωμένες μεταφορές διασφαλίζονται με τη χρήση δημόσιων/ ιδιωτικών κλειδιών και διαφορετικών μορφών συστημάτων κινήτρων, όπως απόδειξη εργασίας ή απόδειξη συμμετοχής (Satoshi, 2017).
- Επειδή δεν χρησιμοποιούνται μεσάζοντες, οι μεταφορές κρυπτονομισμάτων μεταξύ δύο συναλλασσόμενων μερών είναι πιο γρήγορες σε σύγκριση με τις τυπικές μεταφορές χρημάτων. Τα έκτακτα δάνεια στην αποκεντρωμένη χρηματοδότηση είναι ένα καλό παράδειγμα τέτοιων αποκεντρωμένων μεταφορών. Αυτά τα δάνεια, τα οποία

διεκπεραιώνονται χωρίς εγγύηση, μπορούν να εκτελεστούν μέσα σε δευτερόλεπτα και χρησιμοποιούνται στις συναλλαγές (Hamacher, 2021).

- Οι επενδύσεις σε κρυπτονομίσματα μπορούν να αποφέρουν κέρδη. Οι αγορές κρυπτονομισμάτων έχουν εκτοξευθεί σε αξία την τελευταία δεκαετία, φτάνοντας κάποια στιγμή σχεδόν τα 2 τρισεκατομμύρια δολάρια. Από τον Μάιο του 2022, το Bitcoin αποτιμήθηκε σε περισσότερα από 550 δισεκατομμύρια δολάρια στις αγορές κρυπτογράφησης (www.coinmarketcap.com, n.d.).
- Η οικονομία των εμβασμάτων δοκιμάζει μια από τις πιο εμφανείς περιπτώσεις χρήσης κρυπτονομισμάτων. Επί του παρόντος, κρυπτονομίσματα όπως το Bitcoin χρησιμεύουν ως ενδιάμεσα νομίσματα για τον εξορθολογισμό των διασυνοριακών μεταφορών χρημάτων. Έτσι, ένα νόμισμα fiat μετατρέπεται σε Bitcoin (ή άλλο κρυπτονόμισμα), μεταφέρεται εκτός συνόρων και, στη συνέχεια, μετατρέπεται στο νόμισμα fiat προορισμού. Αυτή η μέθοδος απλοποιεί τη διαδικασία μεταφοράς χρημάτων και την καθιστά φθηνότερη.

Μειονεκτήματα

- Αν και ισχυρίζονται ότι είναι μια ανώνυμη μορφή συναλλαγής, τα κρυπτονομίσματα είναι στην πραγματικότητα ψευδώνυμα. Αφήνουν ένα ψηφιακό ίχνος που μπορούν να αποκρυπτογραφήσουν υπηρεσίες όπως το Ομοσπονδιακό Γραφείο Ερευνών (FBI). Αυτό ανοίγει τις δυνατότητες των κυβερνήσεων ή των ομοσπονδιακών αρχών να παρακολουθούν τις οικονομικές συναλλαγές των απλών πολιτών (Perlroth et al., 2021).
- Τα κρυπτονομίσματα έχουν γίνει ένα δημοφιλές εργαλείο με τους εγκληματίες για κακόβουλες δραστηριότητες όπως το ξέπλυμα χρήματος και οι παράνομες αγορές. Γνωστή είναι για παράδειγμα η περίπτωση του Dread Pirate Roberts, ο οποίος διατηρούσε μια αγορά για να πουλάει ναρκωτικά στο dark web. Τα κρυπτονομίσματα έχουν γίνει επίσης αγαπημένη ενασχόληση των χάκερ που τα χρησιμοποιούν για δραστηριότητες ransomware (Myre, 2021).
- Σε αντίθεση με την παραδοσιακή χρηματοδότηση, δεν υπάρχει τρόπος να αντιστρέψετε ή να ακυρώσετε μια συναλλαγή κρυπτονομίσματος αφού έχει ήδη αποσταλεί. Σύμφωνα με ορισμένες εκτιμήσεις, περίπου το ένα πέμπτο όλων των bitcoin δεν είναι πλέον προσβάσιμα λόγω απώλειας κωδικών πρόσβασης ή λανθασμένων διευθύνσεων αποστολής (Frankenfield, 2022).

- Θεωρητικά, τα κρυπτονομίσματα προορίζονται να είναι αποκεντρωμένα, με τον πλούτο τους να διανέμεται μεταξύ πολλών μερών σε ένα blockchain. Στην πραγματικότητα, η ιδιοκτησία είναι πολύ συγκεντρωμένη. Για παράδειγμα, μια μελέτη του MIT διαπίστωσε ότι μόλις 11.000 επενδυτές κατείχαν περίπου το 45% της αυξανόμενης αξίας του Bitcoin (Makarov & Schoar, 2021).
- Ένα από τα δυνατά σημεία των κρυπτονομισμάτων είναι ότι ο καθένας μπορεί να τα εξορύξει χρησιμοποιώντας έναν υπολογιστή με σύνδεση στο Διαδίκτυο. Ωστόσο, η εξόρυξη δημοφιλών κρυπτονομισμάτων απαιτεί σημαντική ενέργεια, μερικές φορές τόση ενέργεια όση καταναλώνουν ολόκληρες χώρες. Το ακριβό ενεργειακό κόστος σε συνδυασμό με το απρόβλεπτο της εξόρυξης έχει συγκεντρώσει την εξόρυξη σε μεγάλες εταιρείες των οποίων τα έσοδα ανέρχονται σε δισεκατομμύρια δολάρια. Σύμφωνα με μελέτη του MIT, το 10% των ατόμων/ εταιρειών αντιπροσωπεύει το 90% της εξορυκτικής του ικανότητας (Makarov & Schoar, 2021).
- Αν και τα blockchain κρυπτονομισμάτων είναι εξαιρετικά ασφαλή, άλλα αποθετήρια κρυπτονομισμάτων, όπως ανταλλακτήρια και πορτοφόλια, μπορούν να παραβιαστούν. Πολλά ανταλλακτήρια κρυπτονομισμάτων και πορτοφόλια έχουν χακαριστεί, με αποτέλεσμα μερικές φορές να κλαπούν «κέρματα» αξίας εκατομμυρίων δολαρίων (Redman, 2019).
- Τα κρυπτονομίσματα που διαπραγματεύονται σε δημόσιες αγορές υφίστανται σημαντική αστάθεια τιμών.

Τα κρυπτονομίσματα είναι ένα νέο παράδειγμα για την έννοια του χρήματος. Λόγος ύπαρξής τους είναι να εξορθολογίσουν την υπάρχουσα χρηματοοικονομική αρχιτεκτονική για να την καταστήσουν ταχύτερη και φθηνότερη. Η τεχνολογία και η αρχιτεκτονική των κρυπτονομισμάτων βασίζεται στην αποκέντρωση των υφιστάμενων νομισματικών συστημάτων. Επιτρέπουν στα συναλλασσόμενα μέρη να ανταλλάσσουν αξία και χρήμα ανεξάρτητα από ενδιάμεσους οργανισμούς όπως οι τράπεζες. Παρόλ'αυτά η χρήση τους εγκυμονεί και κάποια σημαντικά μειονεκτήματα. Έτσι, ορισμένοι οικονομολόγοι θεωρούν ότι τα κρυπτονομίσματα είναι μια βραχύβια μόδα ή κερδοσκοπική φούσκα.

3.6 Ομοιότητες και διαφορές κρυπτονομισμάτων και ηλεκτρονικού χρήματος

Για την περαιτέρω κατανόηση του όρου κρυπτονομίσματος είναι σκόπιμο να εξεταστούν οι διαφορές και οι ομοιότητες ανάμεσα σε αυτά και το ηλεκτρονικό χρήμα.

Ομοιότητες/ Διαφορές

- Και τα δύο νομίσματα συναντώνται σε ψηφιακή μορφή:
Το ηλεκτρονικό χρήμα είναι ένας τρόπος χρήσης χρημάτων fiat, δηλαδή ένας τρόπος αλληλεπίδρασης με το παραστατικό χρήμα.
- Ανταπόκριση σε διαφορετική λογιστική μονάδα:
Το ηλεκτρονικό χρήμα δεν είναι διαφορετικό νόμισμα από το παραστατικό χρήμα και υπόκειται στα ίδια μέτρα ασφάλειας και ρύθμισης με το εθνικό νόμισμα.
Αντιθέτως, η τιμή - και κατά συνέπεια η αξία - των κρυπτονομισμάτων καθορίζεται από χρήστες που θέλουν να αγοράσουν (ζήτηση) και να πουλήσουν (προσφορά). Συνεπώς διαφέρει ο καθορισμός των τιμών στις οποίες είναι διατεθειμένοι να συναλλάσσονται.
- Διαφορές στην παραγωγή και έκδοση:
Το ηλεκτρονικό χρήμα δημιουργείται μέσω ψηφιακής παραγωγής και εκδίδεται από νομίμως εξουσιοδοτημένο πρόσωπο, όπως μια Κεντρική Τράπεζα. Η δημιουργία ενός νέου κρυπτονομίσματος επηρεάζεται από το περιβάλλον του. Τα κρυπτονομίσματα βασίζονται σε ένα σύστημα που καθορίζει πώς θα δημιουργηθούν, σε ποιον θα τα κατέχουν και τις συνθήκες υπό τις οποίες μπορούν να χρησιμοποιηθούν.
- Ταυτοποίηση χρήστη:
Ενώ τα κρυπτονομίσματα διασφαλίζουν την επιθυμητή ανωνυμία για τους χρήστες, το ηλεκτρονικό χρήμα και η κίνηση που σχετίζεται με τη χρήση του είναι ανιχνεύσιμα. Οι κάτοχοι λογαριασμών κρυπτονομισμάτων έχουν τη δυνατότητα να αποκαλύψουν οικειοθελώς την ταυτότητά τους.

3.7 Ανακεφαλαίωση

Τα κρυπτονομίσματα είναι εικονικά ή με άλλα λόγια ψηφιακά νομίσματα που χρησιμοποιούν ένα ηλεκτρονικό καθολικό με ισχυρή κρυπτογράφηση, έννοια από την οποία έλκουν και το όνομά τους. Για τη διαδικασία της κρυπτογράφησης ή αλλιώς της κωδικοποίησης των πληροφοριών, τα συγκεκριμένα νομίσματα χρησιμοποιούν το blockchain ως την υποκείμενη τεχνολογία για την άμεση καταγραφή, επαλήθευση και ασφάλεια των δεδομένων και των συναλλαγών.

Το πιο δημοφιλές κρυπτονόμισμα είναι το bitcoin, που κατάφερε να φτάσει, από την αξία του 1 δολαρίου το 2011, στην κορυφή των σχεδόν 69 χιλιάδων δολαρίων τον 11/2021, ενώ πριν από λίγες ημέρες διαπραγματευόταν σε τιμή κατώτερη των 20.000 δολαρίων. Όσον αφορά τις αλυσίδες blockchain η πρωτιά λόγω φήμης ανήκει τόσο στο Bitcoin όσο και στο Ethereum, που αν και εμφανίζει μικρότερη αξία στην αγορά από το πρώτο κατάφερε το 2011 να το ξεπεράσει στον αριθμό συναλλαγών.

Παρακολουθώντας την ιστορική αναδρομή και την πορεία των κρυπτονομισμάτων χαρακτηρίζονται ολοένα και πιο δημοφιλείς εναλλακτικές λύσεις για διαδικτυακές πληρωμές. Πριν όμως οποιαδήποτε μετατροπή πραγματικών δολαρίων, ευρώ ή άλλων παραδοσιακών νομισμάτων σε □ (το σύμβολο για το Bitcoin), θα πρέπει να κανείς να αναλογιστεί τι είναι το κρυπτονόμισμα, ποιοι είναι οι κίνδυνοι από τη χρήση και ποιοι οι τρόποι προστασίας.

Κεφάλαιο 4

Θεωρητικό υπόβαθρο μοντέλων πρόβλεψης

4.1 Εισαγωγή

Τα τελευταία χρόνια, υπάρχει μεγάλο ενδιαφέρον ως προς την εξόρυξη δεδομένων και την μελέτη τους. Ο μεγαλύτερος όγκος δεδομένων δίνεται στη μορφή χρονοσειρών. Μια χρονοσειρά είναι απλώς μια σειρά σημείων δεδομένων που ταξινομούνται χρονικά. Σε μια χρονολογική σειρά, ο χρόνος είναι συχνά η ανεξάρτητη μεταβλητή και ο στόχος είναι συνήθως να γίνει μια πρόβλεψη για το μέλλον (Peixeiro, 2019).

Η ανάλυση χρονοσειρών (TimeSeries: TS) είναι η μελέτη μιας ακολουθίας δεδομένων που συλλέγονται μέσα στο χρόνο. Είναι μια αποτελεσματική προσέγγιση για τη διερεύνηση της διακύμανσης των μεταβλητών και επομένως χρησιμοποιείται ευρέως σε πολλούς τομείς της επιστήμης και της μηχανικής (Rhif et al., 2019).

Στο τέταρτο κεφάλαιο δίνεται έμφαση στην θεωρία της ανάλυσης χρονοσειράς. Αφού σχολιαστούν τα βασικά χαρακτηριστικά μιας χρονοσειράς, και οι στατιστικοί έλεγχοι που θα χρησιμοποιηθούν, επεξηγούνται αναλυτικά τα βήματα της πρόβλεψης. Επιπλέον, αναφέρονται τα σημαντικότερα κριτήρια αξιολόγησης των προβλέψεων. Τέλος, παρουσιάζονται θεωρητικά τα μοντέλα πρόβλεψης που θα χρησιμοποιηθούν στο επόμενο κεφάλαιο.

4.2 Βασικά χαρακτηριστικά και ανάλυση χρονοσειρών

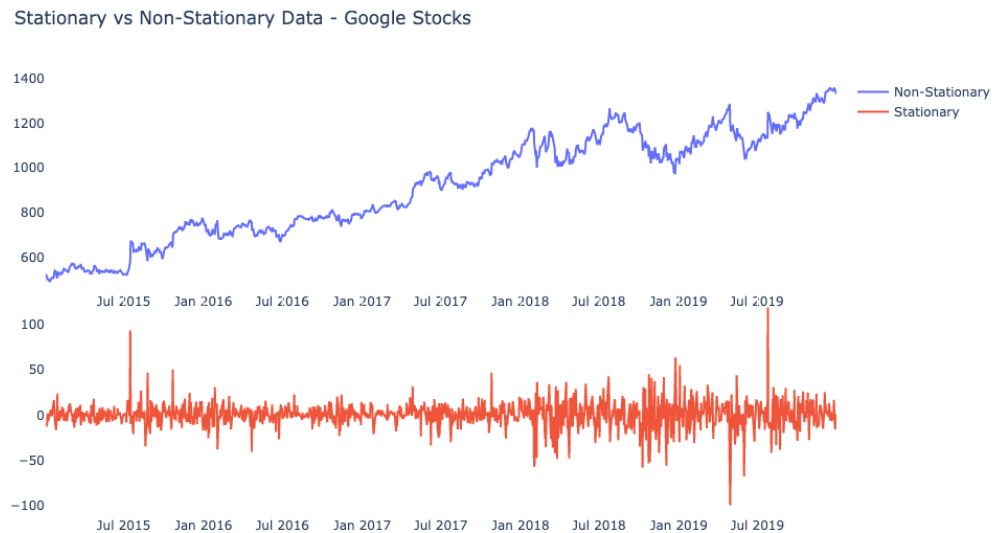
Ποιοτικά χαρακτηριστικά των χρονοσειρών

Όλες οι χρονοσειρές ανεξαρτήτως περιγραφόμενου μεγέθους (τιμή κλεισίματος μετοχής, πωλήσεις αντικειμένων κτλ.) παρουσιάζουν ορισμένα βασικά χαρακτηριστικά τα οποία είναι τα εξής:

1. Στασιμότητα (Stationary):

Η στασιμότητα μίας χρονοσειράς επηρεάζεται από τη συνδιακύμανση στη πάροδο του χρόνου δηλαδή από τη μέση τιμή και την διακύμανση. Εφόσον σε δύο διαδοχικές χρονικές στιγμές t_1 , t_2 τα στατιστικά μεγέθη μίας χρονοσειράς δεν έχουν μεταβληθεί τότε η χρονοσειρά θεωρείται στατική. Για λόγους ευκολίας στην ανάλυση επικρατεί ο έλεγχος της μη μεταβολής της μέσης

τιμής και διακύμανσης. Βιβλιογραφικά γίνεται διάκριση σε αυστηρή και ασθενή στασιμότητα. Ακολουθεί στο διάγραμμα 4.1, ένα παραδείγματα στασιμότητας και μη σε μία χρονοσειρά.



Διάγραμμα 4.1: Παράδειγμα στασιμότητας και μη σε μία χρονοσειρά. Πηγή: <https://blogs.cisco.com/analytics-automation/arima2>

Μαθηματικά, η χρονοσειρά TS (y_t) είναι σταθερή εάν, για όλα τα t ισχύει (Rhif et al., 2019):

Εξίσωση 1: Εξισώσεις- προϋποθέσεις για ύπαρξη σταθερότητας σε χρονοσειρά.

$$E(y_t) = E[(y_{t-1})] = \mu,$$

$$\text{Var}(y_t) = \gamma < \infty,$$

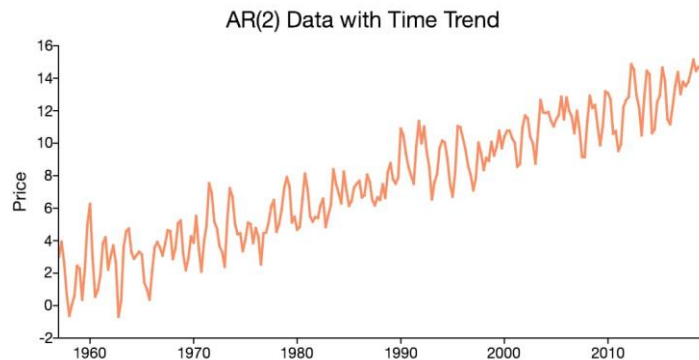
$$\text{Cov}(y_t, y_{t-k}) = \gamma_k$$

Όπου $E()$ είναι η αναμενόμενη τιμή που ορίζεται ως ο μέσος όρος του συνόλου της ποσότητας, $\text{Var}()$ και $\text{Cov}()$ είναι αντίστοιχα, οι συναρτήσεις διακύμανσης και συνδιακύμανσης.

Η σταθερότητα/ στασιμότητα είναι μια σημαντική έννοια στην ανάλυση χρονοσειρών. Για μια συνοπτική εισαγωγή στο θέμα και τους λόγους που την καθιστούν σημαντική αρκεί να πούμε ότι: Α) Μεταφράζεται ως ότι οι στατιστικές ιδιότητες των χρονοσειρών (ή μάλλον η διαδικασία που τη δημιουργεί) δεν αλλάζουν με την πάροδο του χρόνου. Β) Η σταθερότητα είναι σημαντική γιατί πολλά χρήσιμα αναλυτικά εργαλεία και στατιστικές δοκιμές - μοντέλα βασίζονται σε αυτήν (Palachy, 2019). Ως εκ τούτου, η ικανότητα προσδιορισμού του αν μια χρονοσειρά είναι ακίνητη είναι σημαντική.

2. Τάση (Trend):

Η τάση δείχνει τη γενική πορεία των δεδομένων, την ικανότητά τους να αυξάνονται ή να μειώνονται για μεγάλο χρονικό διάστημα. Μια τάση είναι ομαλή, γενική, μακροπρόθεσμη, μέση. Δεν είναι πάντα απαραίτητο η αύξηση ή η μείωση να είναι στην ίδια κατεύθυνση σε όλη τη δεδομένη χρονική περίοδο. Είναι παρατηρήσιμο ότι οι τάσεις μπορεί να αυξηθούν, να μειωθούν ή να είναι σταθερές σε διαφορετικές χρονικές περιόδους. Αλλά η συνολική τάση πρέπει να είναι ανοδική, καθοδική ή σταθερή (toppr.com, 2020). Ο πληθυσμός, η γεωργική παραγωγή, τα προϊόντα που κατασκευάζονται, ο αριθμός γεννήσεων και θανάτων, αριθμός βιομηχανίας ή οποιουδήποτε εργοστασίου, αριθμός σχολείων ή κολεγίων είναι μερικά από τα παράδειγμα του που δείχνουν κάποιου είδους τάσεις μετακίνησης. Η μαθηματική της έκφραση είναι ο κινητός μέσος όρος. Ένα παράδειγμα τάσης προβάλλεται στο διάγραμμα 4.2.



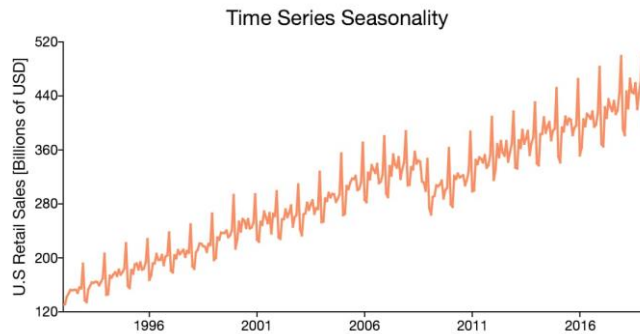
Διάγραμμα 4.2: Παράδειγμα χρονοσειράς με τάση. Πηγή: <https://www.aptech.com/blog/introduction-to-the-fundamentals-of-time-series-data-and-analysis/>

Κάποιες μέθοδοι προσδιορισμού της τάσης είναι η μέθοδος των κινητών μέσων, η μέθοδος της ευθείας ελαχίστων τετραγώνων, η μέθοδος της καμπύλης ελαχίστων τετραγώνων και άλλα.

3. Περιοδικότητα ή Εποχικότητα (Seasonal):

Η εποχικότητα είναι ένα άλλο χαρακτηριστικό των δεδομένων χρονοσειρών που μπορεί να εντοπιστεί οπτικά σε γραφήματα χρονοσειρών (διάγραμμα 4.3). Η εποχικότητα εμφανίζεται όταν τα δεδομένα χρονοσειρών εμφανίζουν κανονικά και προβλέψιμα μοτίβα σε χρονικά διαστήματα μικρότερα από ένα έτος. Η εποχικότητα είναι πάντα μιας σταθερής και γνωστής περιόδου. Ως εκ τούτου, οι εποχιακές χρονοσειρές ονομάζονται μερικές φορές περιοδικές χρονοσειρές (Hyndman, 2011). Τα επαναλαμβανόμενα γεγονότα αποτελούν την εποχικότητα και μπορούν να βοηθήσουν στην άντληση χρήσιμων πληροφοριών. Πολύ

συχνό είναι το φαινόμενο στο χώρο των πωλήσεων. Ένα παράδειγμα χρονοσειράς με εποχικότητα είναι οι λιανικές πωλήσεις, οι οποίες συχνά αυξάνονται μεταξύ Σεπτεμβρίου και Δεκεμβρίου και θα μειωθούν μεταξύ Ιανουαρίου και Φεβρουαρίου.

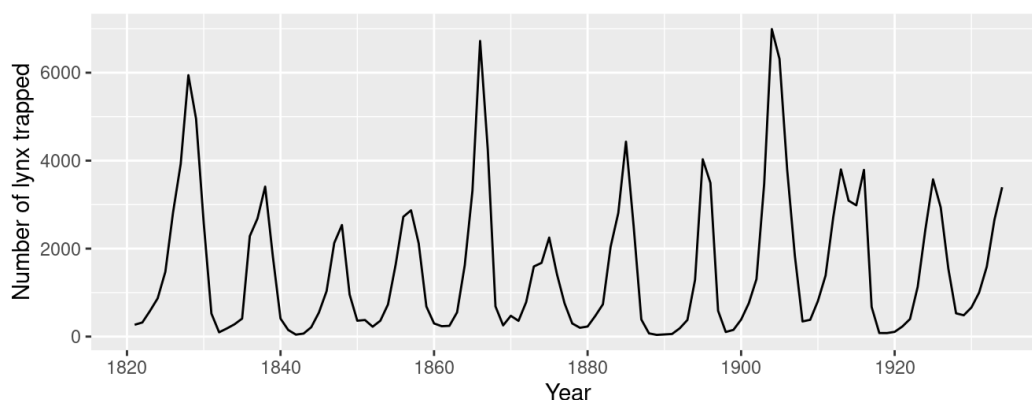


Διάγραμμα 4.3: Παράδειγμα χρονοσειράς με εποχικότητα. Πηγή: <https://www.aptech.com/blog/introduction-to-the-fundamentals-of-time-series-data-and-analysis/>

4. Κυκλικότητα (Cyclical):

Κυκλικό μοτίβο εμφανίζεται όταν τα δεδομένα παρουσιάζουν αυξήσεις και πτώσεις που δεν είναι σταθερής περιόδου. Η διάρκεια αυτών των διακυμάνσεων είναι συνήθως τουλάχιστον 2 χρόνια. Ένα παράδειγμα αποτελούν οι οικονομικοί κύκλοι που συνήθως διαρκούν αρκετά χρόνια, αλλά η διάρκεια του τρέχοντος κύκλου είναι άγνωστη εκ των προτέρων.

Πολλοί άνθρωποι μπερδεύουν την κυκλική συμπεριφορά με την εποχιακή συμπεριφορά, αλλά είναι πραγματικά πολύ διαφορετικές. Αν οι διακυμάνσεις δεν είναι σταθερής περιόδου τότε είναι κυκλικές. Εφόσον η περίοδος είναι αμετάβλητη και σχετίζεται με κάποια πτυχή του ημερολογίου, τότε το μοτίβο είναι εποχιακό. Γενικά, η μέση διάρκεια των κύκλων είναι μεγαλύτερη από τη διάρκεια ενός εποχιακού μοτίβου και το μέγεθος των κύκλων τείνει να είναι πιο μεταβλητό από το μέγεθος των εποχιακών προτύπων (Hyndman, 2011).



Διάγραμμα 4.4: Παράδειγμα χρονοσειράς με κυκλικότητα. Πηγή: <https://robjhyndman.com/hyndsight/cyclists/>

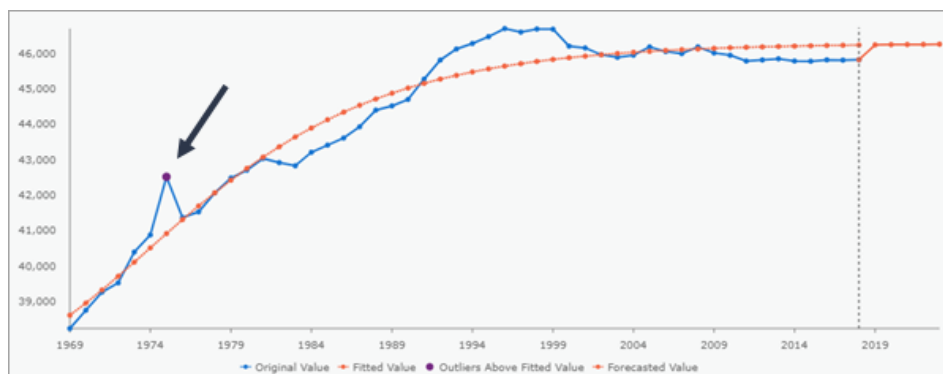
Το διάγραμμα 4.4 απεικονίζει δεδομένα για τον καναδικό λύγκα, συγκεκριμένα τον αριθμό των λύγκα που παγιδεύονται κάθε χρόνο στην περιοχή του ποταμού McKenzie στο βορειοδυτικό άκρο του Καναδά (1821-1934). Αυτά δείχνουν σαφείς απεριοδικούς πληθυσμιακούς κύκλους περίπου 10 ετών. Οι κύκλοι δεν είναι σταθερής διάρκειας καθώς κάποιοι διαρκούν 8 ή 9 χρόνια και άλλοι διαρκούν περισσότερο από 10 χρόνια.

5. Ασυνέχειες (Discontinuity), Ασυνήθιστες τιμές (Outliers):

Ασυνέχειες μπορεί να προκύψουν σε περίπτωση αλλαγής μοντέλου δεδομένων ή ακόμη και παραλλαγών της πηγής σήματος και του περιβάλλοντος. Μια χρονοσειρά δεδομένων μπορεί να ερμηνευθεί ως μια στοχαστική διαδικασία συν μια συνάρτηση βήματος που αντιπροσωπεύει τις ασυνέχειες ή τα άλματα χρονοσειρών (Roggero, 2012). Ο στόχος της ανάλυσης χρονοσειρών είναι η διάκριση μεταξύ στοχαστικών και ντετερμινιστικών σημάτων, τα οποία παράγονται από διαφορετικές πηγές και αναμειγνύονται στις χρονοσειρές δεδομένων. Πριν από την ανάλυση της μακροπρόθεσμης γραμμικής τάσης και των περιοδικών επιδράσεων, είναι απαραίτητο να ανιχνευθούν και να αφαιρεθούν οι ασυνέχειες χρονοσειρών, συχνά μη τεκμηριωμένες.

Οι ασυνήθιστες, ακραίες τιμές στα δεδομένα χρονοσειρών είναι τιμές που διαφέρουν σημαντικά από τα μοτίβα και τις τάσεις των άλλων τιμών στις χρονοσειρές. Για παράδειγμα, μεγάλοι αριθμοί ηλεκτρονικών αγορών κατά τη διάρκεια των εορτών ή μεγάλοι αριθμοί τροχαίων ατυχημάτων κατά τη διάρκεια ισχυρών βροχοπτώσεων μπορεί να ανιχνευθούν ως ακραίες τιμές στις χρονοσειρές τους. Τα απλά σφάλματα εισαγωγής δεδομένων, όπως η παράλειψη του δεκαδικού αριθμού, είναι μια άλλη κοινή πηγή ακραίων τιμών. Ο

προσδιορισμός των ακραίων τιμών στην πρόβλεψη χρονοσειρών είναι σημαντικός επειδή τα ακραία στοιχεία επηρεάζουν το μοντέλο πρόβλεψης που χρησιμοποιείται για την πρόβλεψη μελλοντικών τιμών. Ακόμη και ένας μικρός αριθμός ακραίων τιμών στη χρονοσειρά μιας τοποθεσίας μπορεί να μειώσει την ακρίβεια και την αξιοπιστία των προβλέψεων. Τοποθεσίες με ακραίες τιμές, ιδιαίτερα ακραίες τιμές προς την αρχή ή το τέλος της χρονοσειράς, ενδέχεται να παράγουν παραπλανητικές προβλέψεις.



Διάγραμμα 4.5: Παράδειγμα ακραίας/ ασυνήθιστης τιμής σε χρονοσειρά. Πηγή: <https://pro.arcgis.com/en/pro-app/latest/tool-reference/space-time-pattern-mining/understanding-outliers-in-time-series-analysis.htm>

6. Τυχαιότητα (Randomization):

Τυχαία ή άρρυθμος κύμανση θεωρείται οποιαδήποτε επίδραση στη διαμόρφωση της τιμής της μεταβλητής η οποία δεν οφείλεται σε κάποια από τις παραπάνω συνιστώσες. Πρέπει να τονιστεί, ότι η τυχαία συνιστώσα εμφανίζεται ακανόνιστα με επιδράσεις, που άλλοτε είναι θετικές και άλλοτε αρνητικές. Οι τυχαίες κυμάνσεις οφείλονται σε μη συστηματικές, απρόβλεπτες επιδράσεις. Κάποια παραδείγματα τέτοιων επιδράσεων είναι απρόβλεπτες αλλαγές τιμών στη διεθνή αγορά, ξαφνικές εξαγγελίες κυβερνητικών μέτρων, ασυνήθιστες κλιματολογικές συνθήκες, φυσικές καταστροφές, πολιτικές κρίσεις, κ.λ.π.

Στο σημείο αυτό αξίζει να γίνει ειδική μνεία για δύο χαρακτηριστικές περιπτώσεις χρονοσειράς, τον λευκό θόρυβο και τον τυχαίο περίπατο.

Άλλα χαρακτηριστικά

- Λευκός Θόρυβος (White noise):

Ο λευκός θόρυβος είναι ένας τύπος χρονοσειράς που δεν έχει τάση ή περιοδικότητα. Τέτοιες χρονοσειρές είναι εντελώς τυχαίες και δεν περιέχουν αυτοσυσχέτιση μεταξύ των παρατηρήσεων. Μια χρονοσειρά ονομάζεται λευκός θόρυβος εάν η κατανομή της έχει μέση

τιμή 0. Εάν τα στοιχεία μιας χρονοσειράς λευκού θορύβου ακολουθούν κανονική (Gaussian) κατανομή, τότε η χρονοσειρά ονομάζεται Gaussian white noise.

- Τυχαίος Περίπατος (Random Walk):

Ένας τυχαίος περίπατος ορίζεται ως μια μη στάσιμη χρονοσειρά στην οποία κάθε τυχαία μεταβλητή τη χρονική στιγμή t προκύπτει από την προηγούμενη τυχαία παρατήρηση τη χρονική στιγμή $t-1$. Οι προφανείς συσχετίσεις που φαίνονται σε μια γραφική παράσταση τυχαίου ιστορικού περιπάτου οφείλονται στο ότι το τυχαίο βήμα κάθε φορά βρίσκεται σε ένα σημείο με γνωστό σημείο εκκίνησης.

Προκειμένου να αναπτυχθεί ένα οικονομετρικό μοντέλο που μπορεί να προβλέψει την εξέλιξη μιας από τις παρατηρήσεις σε μια χρονοσειρά, είναι απαραίτητο να ελέγξουμε εάν η χρονοσειρά αποτελείται από λευκό θόρυβο ή τυχαίους περιπάτους. Και στις δύο περιπτώσεις, θα είναι αδύνατο να κάνουμε μια πρόβλεψη γιατί οι χρονοσειρές θα είναι εντελώς τυχαίες και χωρίς μοτίβο, όπως συμβαίνει με τον λευκό θόρυβο, ή μη στάσιμες, όπως συμβαίνει με τον τυχαίο περίπατο. Σε αυτό το κεφάλαιο, θα συζητήσουμε τον τρόπο με τον οποίο μια χρονοσειρά ελέγχεται για σταθερότητα.

Η ανάλυση των χρονοσειράς περιλαμβάνει τρία σημαντικά ειδικά προβλήματα:

1. Πρόβλεψη (Prediction)
2. Μοντελοποίηση (Modeling)
3. Χαρακτηρισμός (Characterization)

Στόχος της πρόβλεψης είναι να υπολογίσει με όσο το δυνατόν μεγαλύτερη ακρίβεια την εξέλιξη του συστήματος για μικρή χρονική διάρκεια, δηλαδή ουσιαστικά να υπολογίσει τις μελλοντικές τιμές της μεταβλητής της χρονοσειράς.

Στόχος της μοντελοποίησης είναι να 'κατανοήσει' τη συμπεριφορά και τις χαρακτηριστικές ιδιότητες του συστήματος για μεγάλη χρονική διάρκεια, μέσω της διαδοχής των μετρήσεων.

Στόχος του χαρακτηρισμού του συστήματος είναι ο καθορισμός κάποιων θεμελιωδών ιδιοτήτων του συστήματος.

4.3 Ορισμός πρόβλεψης χρονοσειράς

Η ανάλυση χρονοσειρών μπορεί να είναι πολύ χρήσιμη για την παρατήρηση της εξέλιξης των μεγεθών που περιγράφονται από τα δεδομένα, καθώς και για την εκτίμηση της μελλοντικής πορείας της ακολουθίας των παρατηρήσεων. Η πρόβλεψη της μελλοντικής τιμής μιας μεταβλητής βασίζεται στην προηγούμενη συμπεριφορά της. Τα προγνωστικά μοντέλα έχουν βρεθεί ότι είναι επιτυχή στην τυποποίηση αυτής της διαδικασίας. Η πρόβλεψη έχει γίνει όλο και πιο σημαντική τα τελευταία χρόνια, με πολλούς ανθρώπους να τις αναζητούν για καθοδήγηση σχετικά με μελλοντικά γεγονότα. Ανάλογα με τη σειρά των γεγονότων, οι προσεγγίσεις για την επίλυση ενός προβλήματος μπορεί να είναι διαφορετικές.

Στο αντικείμενο των προβλέψεων κύριος στόχος είναι η κατά το δυνατόν μεγαλύτερη ευστοχία και η ελαχιστοποίηση της απόκλισής τους από τις πραγματικές μελλοντικές τιμές. Με αυτόν τον τρόπο, οι μέθοδοι πρόβλεψης συμβάλλουν στη σωστή και έγκαιρη λήψη αποφάσεων.

Η πρόβλεψη μέσω των χρονοσειρών χρησιμοποιείται ευρέως και συχνά σε πλήθος επιστημονικών κλάδων. Τα δεδομένα, λοιπόν, περιέχουν όλη τη πληροφορία αλλά χρειάζεται και η κριτική ικανότητα του επιστήμονα, για να επιτευχθούν ικανοποιητικά ακριβείς προβλέψεις. Στην πράξη, εγκυμονούν παράγοντες που εισάγουν σφάλματα στην παραπάνω διαδικασία. Ένας τέτοιος παράγοντας είναι η αβεβαιότητα. Η αβεβαιότητα αναφέρεται σε απρόσμενα γεγονότα, που επηρεάζουν την πραγματική μελλοντική πορεία της χρονοσειράς υπό εξέταση, και έχει σαν αποτέλεσμα τα διάφορα μοντέλα να εμφανίζουν από μικρή απόκλιση, έως και ολική αστοχία. Επίσης, ένας επιπλέον παράγοντας είναι ότι τα δεδομένα μπορεί να μην είναι αρκετά σε αριθμό για τη διενέργεια μιας σωστής εκτίμησης. Σε τέτοιες περιπτώσεις, η εμπειρία, η ευρηματικότητα και η κρίση του επιστήμονα θα προβάλλει λύσεις.

4.4 Βήματα πρόβλεψης

Καθώς οι τεχνικές πρόβλεψης πηγάζουν από δεδομένα προερχόμενα από ιστορικά γεγονότα, προσδιορίζονται τα πέντε ακόλουθα βήματα:

1. Διατύπωση προβλήματος, συλλογή και προεπεξεργασία των δεδομένων.
2. Τελική επιλογή μέσω διαχείρισης δεδομένων. Απόφαση των m τιμών lag (καθυστερήσεων).

3. Κατασκευή υποδείγματος. Διαχωρισμός των δεδομένων παρατήρησης σε δεδομένα εκπαίδευσης και δεδομένα δοκιμής (ελέγχου).
4. Εφαρμογή υποδείγματος. Χρήση των δεδομένων για εκπαίδευση του predictor ανάλογα με τον αλγόριθμο εκπαίδευσης που χρησιμοποιείται (learning algorithm) (π.χ. Back Propagation).
5. Υπολογισμός της απόδοσης του εκπαιδευμένου predictor με τα δεδομένα ελέγχου – Αξιολόγηση της πρόβλεψης.

Αρχικά, καθώς πρόκειται για αλληλένδετες ενέργειες, η διατύπωση του προβλήματος και η συλλογή των δεδομένων αντιμετωπίζονται ως ένα ενιαίο βήμα. Πώς όμως γίνεται να προσδιοριστούν τα σωστά δεδομένα που θα χρησιμοποιηθούν; Εάν χρησιμοποιηθεί μια μεθοδολογία ποσοτικής πρόβλεψης, πρέπει να υπάρχει βεβαίωση ότι τα δεδομένα είναι ακριβή. Εάν δεν υπάρχουν διαθέσιμα κατάλληλα δεδομένα, το πρόβλημα μπορεί να χρειαστεί να επαναπροσδιοριστεί. Η συλλογή και η χρήση ακριβών δεδομένων μπορεί να είναι μια πρόκληση στη δημιουργία προβλέψεων.

Στο δεύτερο βήμα, η τελική διαχείριση και σωστή επιλογή των δεδομένων είναι απαραίτητη. Κατά τη διαδικασία πρόβλεψης, ενδέχεται να υπάρχουν πάρα πολλά ή πολύ λίγα δεδομένα. Ορισμένα δεδομένα ενδέχεται να μην σχετίζονται με το πρόβλημα. Σε άλλες περιπτώσεις μπορεί κάποια δεδομένα να είναι κατάλληλα αλλά σε αξιολόγηση συγκεκριμένης περιόδου. Συνήθως απαιτείται κάποια προσπάθεια για τη μετατροπή των δεδομένων στην κατάλληλη μορφή για χρήση σε συγκεκριμένες διαδικασίες πρόβλεψης.

Το τρίτο βήμα στη διαδικασία πρόβλεψης είναι η κατασκευή του μοντέλου, το οποίο είναι ένα εργαλείο πρόβλεψης που είναι κατάλληλο για την ελαχιστοποίηση των σφαλμάτων εφαρμογής. Στη συνέχεια, το μοντέλο αξιολογείται για να προσδιοριστεί η αποτελεσματικότητά του στην πρόβλεψη. Ένα απλό μοντέλο είναι καλύτερο για την αποδοχή των προβλέψεων, καθώς είναι ευκολότερο στην κατανόηση και επομένως στην εμπιστοσύνη. Σαφώς σε αυτό το στάδιο επηρεάζει άμεσα την επιλογή η ανθρώπινη κρίση.

Το τέταρτο βήμα για τη δημιουργία ενός μοντέλου πρόβλεψης είναι να χρησιμοποιηθούν προβλέψεις για να γίνει έλεγχος της θεωρίας. Μπορούν να παραχθούν προβλέψεις μοντελοποιώντας την πραγματικότητα, αφού συλλεχθούν τα κατάλληλα δεδομένα και παρθεί η τελική απόφαση. Πρόσφατα ιστορικά δεδομένα μπορούν να χρησιμοποιηθούν για τον έλεγχο της ακρίβειας της διαδικασίας.

Το πέμπτο και τελευταίο βήμα στην πρόβλεψη είναι η αξιολόγηση των τιμών έναντι των προηγούμενων τιμών. Στη συνέχεια, αυτές οι προβλέψεις συγκρίνονται με γνωστές ιστορικές αξίες και εξετάζονται τυχόν σφάλματα. Αυτό γίνεται για να καθοριστεί πόσο ακριβείς είναι οι προβλέψεις. Η ανάλυση σφαλμάτων συχνά οδηγεί τους αναλυτές να αλλάξουν το μοντέλο πρόβλεψης. Το υποκεφάλαιο 4.4 εξετάζει συγκεκριμένες μεθόδους μέτρησης σφαλμάτων.

4.5 Βασικοί στατιστικοί έλεγχοι

Το πρώτο βήμα για τον έλεγχο της σταθερότητας ή μη μιας χρονοσειράς είναι η οπτική επιθεώρηση της πλοκής της. Όπως ήδη αναφέρθηκε, μια σταθερή χρονοσειρά χαρακτηρίζεται από σταθερό μέσο όρο και σταθερή διακύμανση.

Εκτός από τον οπτικο-εμπειρικό κανόνα, υπάρχουν και στατιστικά τεστ που μπορούν να επιβεβαιώσουν τη στασιμότητα ή τη μη σταθερότητα μιας χρονοσειράς. Τα πιο συχνά χρησιμοποιούμενα τεστ είναι το Augmented Dickey-Fuller Test (ADF) και το Kwiatkowski Phillips Schmidt Shin Test (KPSS). Αυτές οι δύο δοκιμές εξετάζουν εάν υπάρχει μοναδιαία ρίζα στη χρονοσειρά, η οποία θα μπορούσε να υποδεικνύει ότι η σειρά δεν είναι ακίνητη ή ακίνητη.

Augmented Dickey-Fuller Test (ADF):

- Μηδενική υπόθεση H_0 : την ύπαρξη μοναδιαίας ρίζας, δηλαδή μη στάσιμη χρονοσειρά.
- Εναλλακτική υπόθεση H_1 : την μη ύπαρξη μοναδιαίας ρίζας, δηλαδή στάσιμη χρονοσειρά.

Έχει βρεθεί ότι η ασυμπτωτική ισχύς του τεστ σχετίζεται θετικά με το μέγεθος του δείγματος και αρνητικά με τη σειρά της αυτοπαλίνδρομης (Lopez, 1997). Ωστόσο, δεν σχετίζεται με παραμέτρους όχλησης. Χρησιμοποιείται για τον έλεγχο της μηδενικής υπόθεσης ότι η σειρά είναι μη στάσιμη. Όσο πιο αρνητικός είναι ο αριθμός, τόσο πιο εμφανής είναι η απόρριψη της μηδενικής υπόθεσης ότι η χρονοσειρά είναι μη στάσιμη. Επιπλέον, εάν η τιμή p είναι μικρότερη από < 0.05 , μπορούμε να απορρίψουμε τη μηδενική υπόθεση.

Kwiatkowski Phillips Schmidt Shin Test (KPSS):

- Μηδενική υπόθεση H_0 : την μη ύπαρξη μοναδιαίας ρίζας, δηλαδή στάσιμη χρονοσειρά.
- Εναλλακτική υπόθεση H_1 : την ύπαρξη μοναδιαίας ρίζας, δηλαδή μη στάσιμη χρονοσειρά.

Όσο υψηλότερο είναι το στατιστικό της δοκιμής, τόσο πιο εμφανής είναι η απόρριψη της μηδενικής υπόθεσης ότι η σειρά είναι ακίνητη (Khedhiri & Montasser, 2010). Επιπλέον, εάν η τιμή p είναι μικρότερη από < 0.05 , μπορούμε να απορρίψουμε τη μηδενική υπόθεση.

Μόλις επιβεβαιωθεί η σταθερότητα των δεδομένων, ο ερευνητής μπορεί να συνεχίσει τη μελέτη ως συνήθως. Ωστόσο, εάν η χρονική σειρά θεωρείται μη στάσιμη, θα πρέπει να μετατραπεί σε σταθερή προκειμένου να αναπτυχθεί ένα μαθηματικό μοντέλο ικανό να προβλέπει τις μελλοντικές της τιμές με τη μεγαλύτερη δυνατή ακρίβεια.

Οι καθυστερήσεις χρονοσειρών δεν είναι στάσιμες επειδή ο μέσος όρος και η διακύμανση των παρατηρήσεων αλλάζουν με την πάροδο του χρόνου. Τα στοιχεία δείχνουν μια γενική τάση ή εποχικότητα. Η εποχικότητα είναι ένα σύννηθες φαινόμενο σε χρονοσειρές που σχετίζονται με πωλήσεις προϊόντων ή υπηρεσιών, όπως οι αυξήσεις στις πωλήσεις εισιτηρίων κατά τις περιόδους διακοπών.

Για να μετατραπεί μια μη στάσιμη χρονοσειρά σε σταθερή, πρέπει να αφαιρεθεί η τάση και η εποχικότητα. Αυτή η εξάλειψη γίνεται για να «απομονώσουμε» το σύστημα που χρειάζεται να μελετηθεί από τους εξωτερικούς παράγοντες. Στο στάδιο της πρόβλεψης χρονοσειρών, η τάση και η περιοδικότητα προστίθενται στο εκτιμώμενο μοντέλο προκειμένου να παραχθεί η καλύτερη προσαρμογή του μοντέλου στα δεδομένα χρονοσειρών.

4.6 Κριτήρια αξιολόγησης της πρόβλεψης

Για την αξιολόγηση της εφαρμογής των μοντέλων πρόβλεψης στις πραγματικές τιμές τις χρονοσειράς χρησιμοποιούνται κάποιοι σημαντικοί δείκτες. Παρακάτω θα αναλυθούν συντόμως οι πιο σημαντικοί δείκτες και ο τρόπος υπολογισμού τους.

- Μέσο Τετραγωνικό Σφάλμα - Mean Squared Error (MSE)

Το μέσο τετραγωνικό σφάλμα είναι ένας στατιστικός δείκτης που λειτουργεί ως μέτρο της ακρίβειας της πρόβλεψης, δίνοντας όμως περισσότερο βάρος στα μεγάλα σφάλματα και λιγότερο στα μικρά (λογικό αφού ο τύπος είναι υψωμένος στο τετράγωνο). Ο συγκεκριμένος δείκτης χρησιμοποιείται κυρίως ως κριτήριο επιλογής των κατάλληλων παραμέτρων εξομάλυνσης και υπολογίζεται από τον ακόλουθο τύπο:

Εξίσωση 1: Εξίσωση υπολογισμού δείκτη MSE.

$$MSE = \frac{1}{n} \sum_{k=1}^n (x(k) - \hat{x}(k))^2$$

- Ρίζα Μέσου Τετραγωνικού Σφάλματος - Root Mean Squared Error (RMSE)

Έχει όμοιες ιδιότητες με τον δείκτη MSE, μόνο που είναι εκφρασμένος σε μονάδες της αρχικής χρονοσειράς. Ο τύπος που περιγράφει αυτό το στατιστικό δείκτη είναι ο παρακάτω:

Εξίσωση 2: Εξίσωση υπολογισμού δείκτη RMSE.

$$RMSE = \sqrt{\frac{1}{n} \sum_{k=1}^n (x(k) - \hat{x}(k))^2}$$

Από του δείκτες που παρουσιάστηκαν σε αυτό το κεφάλαιο, σημαντικότερος για τον πρόβλημα πρόβλεψης χρονοσειράς, μπορεί να θεωρηθεί ο δείκτης της Ρίζας Μέσου Τετραγωνικού Σφάλματος. Η άποψη αυτή στηρίζεται στο γεγονός πως η πλειοψηφία των ερευνητών που ασχολήθηκαν με το συγκεκριμένο θέμα, έθεσαν ως μέτρα αξιολόγησης αυτούς τους δύο στατιστικούς δείκτες, γεγονός που οδηγεί και την παρούσα εργασία να κινηθεί προς αυτήν την κατεύθυνση.

- Μέσο Απόλυτο Σφάλμα – Absolute Mean Error (AME or MAE)

Ο συγκεκριμένος δείκτης εκφράζει ένα μέτρο της ακρίβειας της πρόβλεψης έναντι των πραγματικών τιμών διατηρώντας τις αρχικές μονάδες της χρονοσειράς. Μεγάλη τιμή του δείκτη σημαίνει μικρή ακρίβεια της μεθόδου πρόβλεψης που εφαρμόστηκε. Η μαθηματική διατύπωση αυτού του στατιστικού δείκτη είναι η εξής:

Εξίσωση 3: Εξίσωση υπολογισμού δείκτη AME.

$$AME = \frac{1}{n} \sum_{k=1}^n |x(k) - \hat{x}(k)|$$

Ο συγκεκριμένος δείκτης έχει χρησιμοποιηθεί αρκετά στην βιβλιογραφία και υπάρχουν πολλές επιστημονικές διαφωνίες για την υπεροχή του ή όχι έναντι του δείκτη RMSE. Οι περισσότερες έρευνες καταλήγουν στο συμπέρασμα ότι δεν είναι σκόπιμη η σύγκριση και είναι καλό να συνυπολογίζονται όλοι οι δείκτες (Chai & Draxler, 2014).

- Μέσο Απόλυτο Ποσοστιαίο Σφάλμα - Mean Absolute Percentage Error (MAPE or MRE (%))

Το μέσο απόλυτο ποσοστό σφάλματος (MAPE) υπολογίζεται χρησιμοποιώντας απόλυτα σφάλματα σε κάθε περίοδο διαιρεμένα με τις πραγματικές τιμές παρατήρησης για αυτήν την περίοδο. Στη συνέχεια, υπολογίζεται το μέσο απόλυτο ποσοστό σφάλματος. Αυτή η προσέγγιση είναι χρήσιμη όταν το μέγεθος ή το μέγεθος των προγνωστικών μεταβλητών είναι σημαντικό για την αξιολόγηση της ακρίβειας των προβλέψεων. Συγκεκριμένα, είναι το άθροισμα των μεμονωμένων απόλυτων σφαλμάτων διαιρούμενο με τη ζήτηση (κάθε περίοδο ξεχωριστά) και ο τύπος φαίνεται παρακάτω:

Εξίσωση 4: Εξίσωση υπολογισμού δείκτη MRE(%).

$$MRE(\%) = \frac{100}{n} \sum_{k=1}^n \frac{|x(k) - \hat{x}(k)|}{x(k)}$$

Ένα από τα πλεονεκτήματα που προσφέρει ο συγκεκριμένος δείκτης είναι ότι εκφράζεται ως ποσοστό, που είναι ανεξάρτητο από την κλίμακα. Το MAPE υποδεικνύει πόσο σφάλμα στην πρόβλεψη συγκρίνεται με την πραγματική τιμή της σειράς. Μπορεί επίσης να χρησιμοποιηθεί για να συγκρίνει την ακρίβεια των ίδιων ή διαφορετικών μεθόδων σε δύο διαφορετικές σειρές και να μετρήσει την ακρίβεια της εκτιμώμενης τιμής του μοντέλου εκφρασμένη σε όρους απόλυτου μέσου όρου σφάλματος ποσοστού (Prayudani et al., 2019).

Ωστόσο, το MAPE έχει το σημαντικό μειονέκτημα ότι παράγει άπειρες ή απροσδιόριστες τιμές για μηδενικές ή κοντά στο μηδέν πραγματικές τιμές (Kim & Kim, 2016). Αυτό οφείλεται στο γεγονός ότι το ποσοστό σφάλματος δεν μπορεί να υπερβαίνει το 100% για προβλέψεις που είναι πολύ χαμηλές. Ενώ δεν υπάρχει ανώτερο όριο για τις προβλέψεις που είναι πολύ υψηλές. Συμπερασματικά, το MAPE θα ευνοήσει μοντέλα που κάνουν υπό-πρόβλεψη (under-forecast) παρά αυτά που κάνουν υπέρ-πρόβλεψη (over-forecast).

4.7 Παρουσίαση των προς μελέτη μοντέλων πρόβλεψης

Με την βοήθεια της Python και των βιβλιοθηκών της, θα εφαρμοστούν στα δεδομένα των χρονοσειρών το παραδοσιακό μοντέλο αυτοπαλίνδρομου ολοκληρωμένου κινούμενου μέσου όρου (ARIMA), το επαναλαμβανόμενο νευρωνικό δίκτυο βραχυπρόθεσμης και μακροπρόθεσμης μνήμης (LSTM) και το μοντέλο δομής δένδρου αποφάσεων τυχαίου δάσους

(Random Forest). Τα μοντέλα που επιλέχθηκαν είναι βασικά μοντέλα πρόβλεψης χρηματιστηριακών τιμών, διαφορετικής λογικής και δομής.

Τα παραδοσιακά, ολοκληρωμένα, αυτοπαλινδρομικά μοντέλα κινητών μέσων όρων είναι στοχαστικά μαθηματικά μοντέλα τα οποία βοηθούν στην ανάλυση και πρόβλεψη της εξέλιξης των μεγεθών. Στην γενική τους μορφή, τα συγκεκριμένα μοντέλα περιέχουν τον τυχαίο παράγοντα (σφάλμα πρόβλεψης), τιμές του μεγέθους που εμφανίστηκαν σε προηγούμενες περιόδους και σχετικούς στοχαστικούς παράγοντες. Πιο συγκεκριμένα, κάθε παρόμοιο μοντέλο μπορεί να εκφραστεί ως γραμμικός συνδυασμός των παραπάνω παραγόντων και στόχος μας είναι να ανακαλύψουμε εκείνον που παράγει τις καλύτερες προβλέψεις. Τέλος, η εφαρμογή μοντέλων αυτών προϋποθέτει την εξαγωγή βραχυπρόθεσμων προβλέψεων.

Οι αλγόριθμοι τεχνητών νευρωνικών δικτύων έχουν απασχολήσει έντονα τον χώρο της Τεχνητής Νοημοσύνης. Χρησιμοποιούν έναν επαναλαμβανόμενο βρόχο δεδομένων και πληροφοριών και αναγνώριση προτύπων μέσω ενός καθορισμένου αριθμού κύκλων για να καθορίσουν το προβλεπόμενο αποτέλεσμα. Μία από τις πιο σημαντικές πτυχές ενός Νευρωνικού Δικτύου είναι ότι απαιτούν τεράστιες ποσότητες προηγούμενων περιπτώσεων αποτελεσμάτων, γνωστών ως «δεδομένα εκπαίδευσης», για να βρουν μοτίβα, προκειμένου να γενικεύσουν το δικό του αναμενόμενο αποτέλεσμα. Αν και οι αλγόριθμοι πίσω από την τεχνολογία υπάρχουν εδώ και δεκαετίες, η επεξεργαστική ισχύς που είναι απαραίτητη για την εφαρμογή στην πραγματική ζωή έχει καρποφορήσει μόλις τα τελευταία χρόνια.

4.7.1 Auto-Regressive Integrated Moving Average (ARIMA)

Το μοντέλο ARIMA (Auto-Regressive Integrated Moving Average) χρησιμοποιείται σε δεδομένα χρονοσειρών. Εφευρέθηκε από τους George Box και Gwilym Jenkins και είναι ένα ευρέως χρησιμοποιούμενο μοντέλο πρόβλεψης (Jang & Lee, 2018). Το μοντέλο ARIMA αποτελείται από τους όρους AutoRegressive (AR) και τους όρους Moving Average (MA). Το μοντέλο αυτόματης παλινδρόμησης (AR) και το μοντέλο κινούμενου μέσου όρου (MA), τα οποία είναι μοντέλα παλινδρόμησης που χρησιμοποιούν προηγούμενες τιμές (υστερήσεις) για την πρόβλεψη μελλοντικών τιμών. Για τη θεωρία πίσω από το ARIMA, θα γίνει εξήγηση των δύο συστατικών του χωριστά.

➤ Μοντέλο AR

Επειδή λειτουργούν με βάση την αντίληψη ότι οι προηγούμενες τιμές επηρεάζουν τις τρέχουσες τιμές, τα μοντέλα AR είναι δημοφιλή για την ανάλυση διαδικασιών που μεταβάλλονται χρονικά. Σε ένα μοντέλο AR, μπορούμε να προβλέψουμε τη μεταβλητή που μας ενδιαφέρει χρησιμοποιώντας έναν γραμμικό συνδυασμό των προηγούμενων τιμών της μεταβλητής. Ωστόσο, επειδή βασίζονται αποκλειστικά σε ιστορικά δεδομένα, υποθέτουν αφελώς ότι οι θεμελιώδεις δυνάμεις που επηρέασαν τις προηγούμενες τιμές δεν θα αλλάξουν με την πάροδο του χρόνου, οδηγώντας σε πιθανές απροσδόκητες και ανακριβείς προβλέψεις.

➤ Μοντέλο MA

Αντί να χρησιμοποιεί προηγούμενες τιμές πρόβλεψης σε μια παλινδρόμηση, ένα μοντέλο κινητού μέσου όρου χρησιμοποιεί προηγούμενα σφάλματα πρόβλεψης σε ένα μοντέλο που μοιάζει με παλινδρόμηση. Το μοντέλο βασίζεται στην ιδέα ότι μπορούμε να προσαρμόσουμε το μοντέλο για να βελτιώσουμε τις δυνατότητες πρόβλεψής του, γνωρίζοντας πόσο μακριά ήταν η πρόβλεψή μας από το χθες σε σύγκριση με την πραγματική τιμή.

➤ Όρος I

Το I είναι ένας όρος που προσδιορίζει τη σειρά διαφοροποίησης που πρέπει να εφαρμοστεί στη χρονοσειρά για να γίνει «στάσιμη». Η διαφορά είναι απλώς η αφαίρεση της προηγούμενης τιμής από την τρέχουσα τιμή στη χρονοσειρά. Η τιμή που αντιπροσωπεύει το "I" είναι ο αριθμός των φορών που γίνεται αυτό για να επιτευχθεί σταθερότητα. Μια χρονοσειρά είναι ακίνητη όταν τα χαρακτηριστικά της, όπως ο μέσος όρος και η διακύμανση, δεν επηρεάζονται από το χρόνο μέτρησης. Κατά συνέπεια, μια χρονοσειρά με μακροπρόθεσμα χαρακτηριστικά μοτίβα, δηλαδή κύκλους, τάσεις ή εποχικότητα, δεν είναι στάσιμη. Επομένως, εάν η διαφορά της χρονικής σειράς μία φορά δεν έχει ως αποτέλεσμα σταθερό μέσο όρο και διακύμανση, η διαδικασία θα πρέπει να επαναληφθεί μέχρι να επιτευχθεί αυτό, αν και με προσοχή, και όχι με υπερβολική διαφορά.

Συνεπώς το μοντέλο ARIMA χαρακτηρίζεται από τις τιμές APIMA(p,d,q) όπου:

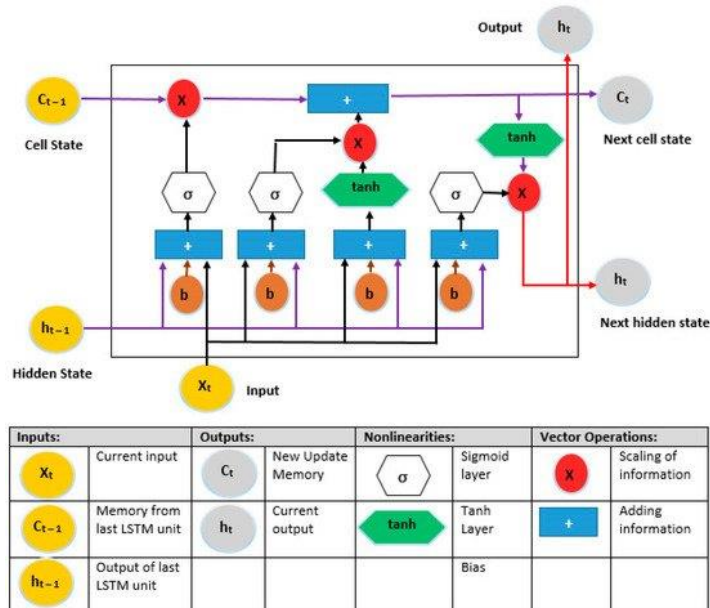
- p είναι ο αριθμός των αυτοπαλινδρομικών όρων,
- d είναι ο αριθμός των μη εποχικών διαφορών που απαιτούνται για σταθερότητα και

- q είναι ο αριθμός των καθυστερημένων σφαλμάτων πρόβλεψης στην εξίσωση πρόβλεψης.

4.7.2 Long Short-Term Memory (LSTM)

Το LSTM (Long Short-Term Memory) πρόκειται για μοντέλο πρόβλεψης με βάση την μνήμη. Είναι ένας τύπος αρχιτεκτονικής επαναλαμβανόμενου νευρωνικού δικτύου (RNN) που έχει χρησιμοποιηθεί με επιτυχία ως επί το πλείστον στην επεξεργασία φυσικής γλώσσας και στην πρόβλεψη χρονοσειρών. Τα LSTM διαφέρουν από τα άλλα νευρωνικά δίκτυα μέσω της χρήσης τους για μάθηση από στοιχεία από την αρχή της σειράς, όχι μόνο από τη σειρά αμέσως προηγούμενη.

Για διάφορα ζητήματα μάθησης που περιλαμβάνουν διαδοχικά δεδομένα, τα επαναλαμβανόμενα νευρωνικά δίκτυα με μακροπρόθεσμη βραχυπρόθεσμη μνήμη (LSTM) έχουν αναδειχθεί ως μια αποτελεσματική και επεκτάσιμη προσέγγιση. Είναι χρήσιμα για την αποτύπωση μακροπρόθεσμων χρονικών εξαρτήσεων, καθώς είναι γενικά και αποτελεσματικά (Greff et al., 2017). Το LSTM είναι μια αρχιτεκτονική τύπου RNN με πύλες που διέπουν τη ροή πληροφοριών μεταξύ των κελιών. Οι δομές πύλης εισόδου και λήθης μπορούν να τροποποιήσουν τις πληροφορίες που ταξιδεύουν κατά μήκος της κατάστασης κυψέλης, με την τελική έξοδο να είναι μια φιλτραρισμένη έκδοση της κατάστασης κελιού με βάση το περιβάλλον από τις εισόδους (Sherratt et al., 2021). Η εικόνα 4.1 απεικονίζει τη δομή ενός αλγορίθμου LSTM (Hamayel & Owda, 2021).



Εικόνα 4.1: Η δομή ενός αλγόριθμου μακράς βραχυπρόθεσμης μνήμης (LSTM).

Στο πλαίσιο μιας πρόβλεψης χρονοσειρών για τα κρυπτονομίσματα, τα LSTM λαμβάνουν υπόψη τα αποτελέσματα τη στιγμή που δημιουργήθηκαν για πρώτη φορά το νόμισμα, καθώς και τα δεδομένα των τελευταίων έξι μηνών, προκειμένου να κάνουν μια πρόβλεψη που έχει μια πιο στρογγυλεμένη κατανόηση του όλη η εικόνα. Λόγω αυτών και ορισμένων άλλων τεχνικών λόγων, κυρίως όσον αφορά τις δυνατότητες βελτιστοποίησης σε ένα LSTM που μειώνουν τα προβλήματα γύρω από τις διαβαθμίσεις που εξαφανίζονται, αυτό τα καθιστά τον ιδανικό υποψήφιο για πρόβλεψη της τιμής των κρυπτονομισμάτων.

4.7.3 Random Forest (RF)

Πρόσφατα αναπτύχθηκε μία νέα κατηγορία μοντέλων παλινδρόμησης για την αντιμετώπιση της πρόβλεψης χρονοσειράς που σχετίζονται με τα κλασικά μοντέλα. Στην παρούσα εργασία θα γίνει χρήση του μοντέλου Random Forest , το οποίο ξεκινά με τη δημιουργία δέντρων αποφάσεων (Breiman, 2001a). Τα δέντρα απόφασης διαχωρίζουν αναδρομικά τα δεδομένα στο χώρο παλινδρόμησης έως ότου το μέγεθος της διακύμανσης στον υποχώρο είναι μικρό. Ένας προγνωστικός δείκτης για τον υποχώρο μπορεί στη συνέχεια να δημιουργηθεί λαμβάνοντας τη μέση τιμή των εξαρτημένων δεδομένων που αντιστοιχούν στα ανεξάρτητα δεδομένα στον υποχώρο. Το αναδρομικό βήμα κατάτμησης μπορεί να απεικονιστεί ως δέντρο. Οι προβλέψεις για νέα δεδομένα λαμβάνονται με την εύρεση του προγνωστικού δείκτη που αντιστοιχεί στο διαμέρισμα όπου βρίσκεται η νέα μεταβλητή εισόδου.

Με την ολοκλήρωση της διαδικασίας δειγματοληψίας εκπαιδεύεται ένα νέο δέντρο. Η συλλογή ή το σύνολο των δέντρων ονομάζεται δάσος. Στη συνέχεια, οι προβλέψεις μπορούν να γίνουν με βάση ένα άθροισμα των μεμονωμένων προβλέψεων που γίνονται από κάθε ένα από τα δέντρα, μια διαδικασία που ονομάζεται «ψηφοφορία». Μια συνολική πρόβλεψη αποκτά καλύτερη ακρίβεια από οποιοδήποτε από τα συστατικά δέντρα. Τα τυχαία δάση ειδικότερα και η προσέγγιση πολλαπλών προβλέψεων γενικά.

Ενώ τα μοντέλα Random Forest συχνά παρέχουν ανώτερη ακρίβεια πρόβλεψης, είναι συχνά δύσκολο να ερμηνευτούν (Breiman, 2001b), αναφέρονται σε «αλγοριθμικά μοντέλα», συμπεριλαμβανομένων των τυχαίων δασών ως «μαύρου κουτιού». Ειδικά για τα Random Forests, κάθε ένα από τα συστατικά δέντρα εκπαιδεύεται σε έναν δυνητικά μη γραμμικό χώρο παλινδρόμησης και στη συνέχεια συνδυάζεται με άλλα. Ως αποτέλεσμα, δεν κωδικοποιούν τα απλά και εύκολα κατανοητά αποτελέσματα που παρέχονται από τα κλασικά γραμμικά μοντέλα. Τα διαγνωστικά για αυτά τα μοντέλα τείνουν να επικεντρώνονται σε μεταβλητή

σημασία. Αυτή η τεχνική αφαιρεί τις μεταβλητές παλινδρόμησης από ένα σύνολο και μετρά την επίδραση στην ακρίβεια της πρόβλεψης. Η προσέγγιση παρέχει ένα λειτουργικό μέτρο της επιρροής που έχει κάθε μεταβλητή στην ακρίβεια χωρίς να παρέχει ένα ερμηνεύσιμο μέτρο για το πώς η μεταβλητή βοηθά στον προσδιορισμό της πρόβλεψης (Kane et al., 2014).

4.8 Ανακεφαλαίωση

Η ανάλυση χρονοσειρών (TimeSeries: TS) είναι η μελέτη μιας ακολουθίας δεδομένων που συλλέγονται μέσα στο χρόνο. Είναι μια αποτελεσματική προσέγγιση για τη διερεύνηση της διακύμανσης των μεταβλητών και επομένως χρησιμοποιείται ευρέως σε πολλούς τομείς της επιστήμης και της μηχανικής. Τα κύρια χαρακτηριστικά που αναλύονται και μελετώνται σε μία χρονοσειρά είναι η: στασιμότητα, τάση, περιοδικότητα, κυκλικότητα, ασυνέχειες, τυχαιότητα, λευκός θόρυβος και τυχαίος περίπατος. Σκοπός της ανάλυσης είναι η δημιουργία προβλέψεων.

Υπάρχουν πολλά μοντέλα πρόβλεψης και βασικοί στατιστικοί έλεγχοι (πχ ADF, KPSS), που βοηθούν στην επιλογή των καλύτερων παραμέτρων του μοντέλου πρόβλεψης για την εφαρμογή του στην προς μελέτη χρονοσειρά. Στην παρούσα εργασία σχολιάζονται διεξοδικά τα μοντέλα ARIMA, RF και LSTM. Για την αξιολόγηση και επιλογή του κατάλληλου μοντέλου χρησιμοποιούνται κάποιοι σημαντικοί δείκτες όπως οι: MAE, MAPE, MSE και RMSE κ.α. Βιβλιογραφικά, φαίνεται να προτιμάται η επιλογή του δείκτη RMSE, όμως απαραίτητη προϋπόθεση για την διεξαγωγή συμπερασμάτων είναι ο συνυπολογισμός περισσότερων κριτηρίων αξιολόγησης.

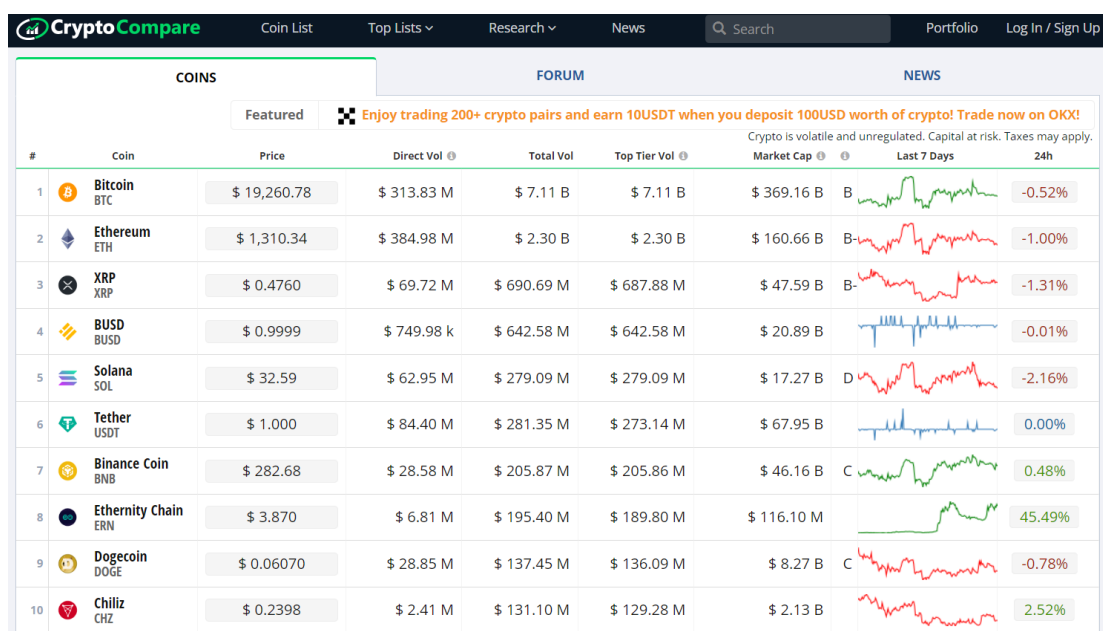
Κεφάλαιο 5

5 Εφαρμογή των μοντέλων πρόβλεψης

5.1 Εισαγωγή

Η εφαρμογή μοντέλων πρόβλεψης για την εκτίμηση των τιμών των κρυπτονομισμάτων έχει προξενήσει μεγάλο ενδιαφέρον σε επενδυτές και ερευνητές. Πρόκειται για πρόκληση καθώς ο όγκος των δεδομένων και η δυναμική μορφή και μεταβολή των τιμών τους είναι αξιοθαύμαστη.

Στην παρούσα εργασία για την δημιουργία μοντέλων πρόβλεψης, τα δεδομένα ημερήσιου όγκου και δείκτη τιμών για τα προς μελέτη κρυπτονομίσματα έναντι του δολαρίου ΗΠΑ είναι από την ημέρα δημιουργίας τους ως τις 26 Σεπτεμβρίου 2022. Μετά την εξαγωγή από το www.cryptocompare.com, η έρευνα για ανάλυσή τους, παρουσίαση και δημιουργία μοντέλων πρόβλεψης διεξάγεται με την βοήθεια της γλώσσας προγραμματισμού Python.



Εικόνα 5.1: Παρουσίαση ιστοσελίδας εξόρυξης δεδομένων χρονοσειράς τιμών κρυπτονομισμάτων (Στιγμιότυπο οθόνης 2/10/2022).

Ερευνητές παγκοσμίως έχουν χρησιμοποιήσει διάφορες τεχνικές και μεθοδολογίες από τους τομείς της στατιστικής, της μηχανικής μάθησης και της βαθιάς μάθησης, όπως LR, ARIMA, LDA, DT, RF, XGBoost, QDA, SVM και LSTM, για να προβλέψουν την τιμή του BTC. Σε

αυτή την ενότητα θα γίνει προσπάθεια πρόβλεψης της τιμής του BTC με τρία μοντέλα: ARIMA, LSTM, RANDOM FOREST καθώς και σύγκριση των αποτελεσμάτων.

5.2 Πρόβλεψη της τιμής του BTC

Πρώτο προς μελέτη κρυπτονόμισμα στην παρούσα έρευνα αποτέλεσε το Bitcoin. Η επιλογή του δεν είναι τυχαία καθώς όπως προαναφέρθηκε καταλαμβάνει την πρώτη θέση στην αγορά και η τιμή του επηρεάζει την τάση του συνόλου των κρυπτονομισμάτων.

5.2.1 Μέθοδος ARIMA

Παρουσιάζονται αναλυτικά τα βήματα και η διαδικασία επιλογής παραμέτρων και εφαρμογής του μοντέλου στην χρονοσειρά.

1. Λήψη δεδομένων

Για να κάνουμε προβλέψεις για το μέλλον, χρειαζόμαστε ιστορικά δεδομένα για να εκπαιδεύσουμε το μοντέλο μας.

2. Ανάλυση δεδομένων σε πλαίσιο δεδομένων και καθαρισμός

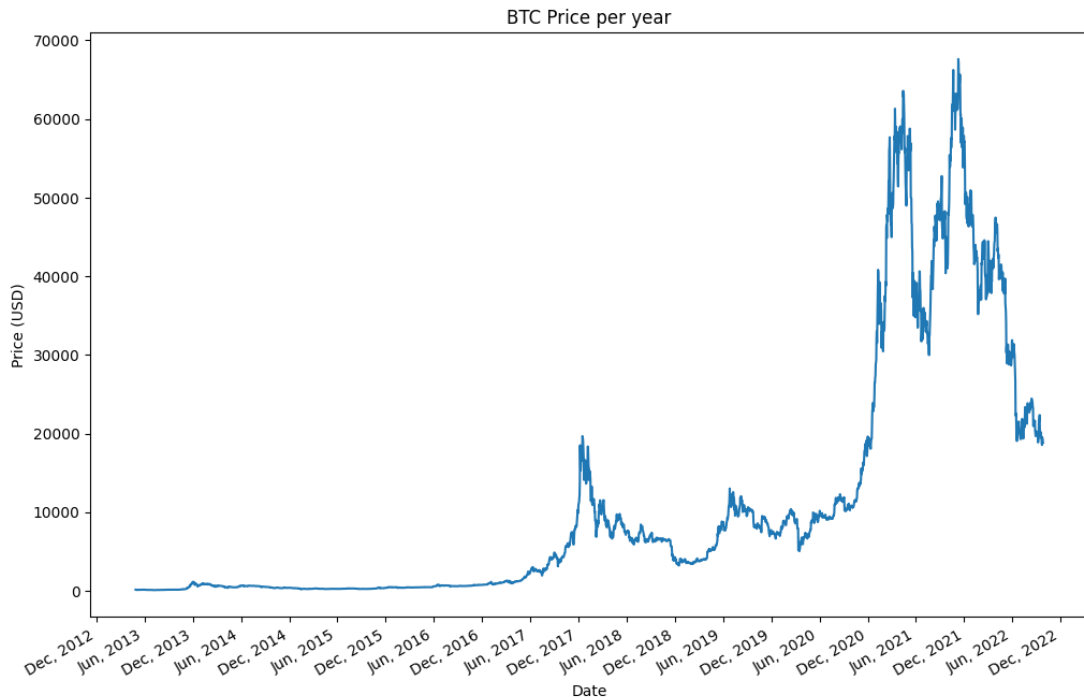
Στη συνέχεια, τα δεδομένα πρέπει να αναλυθούν σε ένα πλαίσιο δεδομένων pandas και να γίνει κάποιος καθαρισμός. Οι στήλες μετονομάζονται για να είναι πιο διαειθητικές και η στήλη ημερομηνίας που είναι σε μορφή ms, θα μετατραπεί σε μορφή ημέρας. Στον πίνακα 5.1 που ακολουθεί φαίνεται μια μικρογραφία του πλαισίου δεδομένων που δημιουργήθηκε.

Πίνακας 5.1: Παρουσίαση τεσσάρων πρώτων ενδεικτικών τιμών του dataframe με αναπαράσταση της ώρας και της τιμής του BTC.

	Date	price
0	2013-04-28 00:00:00+00:00	135.30
1	2013-04-29 00:00:00+00:00	141.96
2	2013-04-30 00:00:00+00:00	135.30
3	2013-05-01 00:00:00+00:00	117.00
4	2013-05-02 00:00:00+00:00	103.43

3. Δοκιμή για Σταθερότητα

Όπως έχει συζητηθεί προηγουμένως, οι χρονοσειρές μας πρέπει να είναι σταθερές. Έτσι, πρώτα, μπορούμε να σχεδιάσουμε τα δεδομένα και να δούμε οπτικά πώς συμπεριφέρονται (διάγραμμα 5.1).



Διάγραμμα 5.1: Αναπαράσταση ιστορικών στοιχείων χρονοσειράς για την τιμή του BTC.

4. Μείωση της Διακύμανσης

Υπάρχει μια σαφής ανοδική τάση, μια μη σταθερή απόκλιση που απεικονίζει μια μη στάσιμη χρονοσειρά. Το πρώτο βήμα για να γίνει η σειρά ακίνητη είναι η σταθεροποίηση της διακύμανσης. Η εκτέλεση ενός μετασχηματισμού καταγραφής στα δεδομένα (log transformation) μπορεί να βοηθήσει σε αυτό όπως φαίνεται και στο διάγραμμα 5.2.



Διάγραμμα 5.2: Εντοπισμός και αφαίρεση τάσης στην χρονοσειρά BTC.

Αν και υπάρχει βελτίωση, η σειρά οπτικά δεν φαίνεται ακόμη να είναι σταθερή και θα χρειαστούν περαιτέρω μετασχηματισμοί για να επιτύχουμε τη σταθερότητα. Ωστόσο, μπορούν να πραγματοποιηθούν δύο προαναφερόμενες δοκιμές για να διαπιστωθεί εάν η σταθερότητα. Αυτές είναι :

- Δοκιμή Augmented Dicky Fuller (Δοκιμή ADF): Η ασυμπτωτική ισχύς της δοκιμής Augmented Dickey Fuller προκύπτει όταν υπάρχουν στοιχεία MA στη διαδικασία δημιουργίας δεδομένων.
- Το Kwiatkowski–Phillips–Schmidt–Shin (KPSS): Χρησιμοποιείται για τον έλεγχο της μηδενικής υπόθεσης ότι η σειρά είναι ακίνητη.

Πίνακας 5.2: Αποτελέσματα ADF και KPSS test μετά την αλλαγή της μεταβλητότητας.

ADF Statistic:	-1.417639
p-value:	0.573791
Results of KPSS Test:	
Test Statistic	0.365008
p-value	0.010000
Lags Used	37.000000
Critical Value(10%)	0.119000
Critical Value(5%)	0.146000
Critical Value(2.5%)	0.176000
Critical Value(1%)	0.216000
dtype:	float64

Τα παραπάνω αποτελέσματα του πίνακα 5.2 δείχνουν ότι το ADF απέτυχε να απορρίψει τη μηδενική υπόθεση, ενώ το KPSS απέρριψε τη μηδενική υπόθεση που σημαίνει ότι η σειρά είναι μη στάσιμη.

Προκειμένου να βρεθεί ο βέλτιστος αριθμός φορών για να διαφοροποιηθούν τα δεδομένα, ορίζεται ένας συνεχής βρόγχος για να συνεχίσει τη διαφορά έως ότου η τιμή p γίνει σημαντική.

Πίνακας 5.3: Αποτελέσματα ελέγχου ADF και KPSS μετά τις διαφοροποιήσεις.

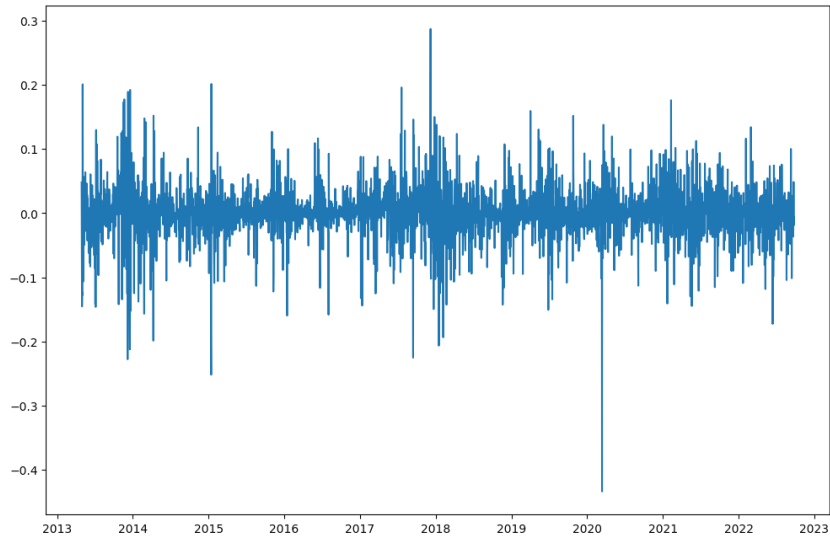
```
ADF Statistic: -15.475431
p-value: 0.000000

Success. Significant values achieved after 1 differencing
Results of KPSS Test:
Test Statistic          0.084328
p-value                 0.100000
Lags Used               16.000000
Critical Value(10%)     0.119000
Critical Value(5%)      0.146000
Critical Value(2.5%)    0.176000
Critical Value(1%)      0.216000
dtype: float64
```

Μετά από μία επανάληψη, ο βρόγχος εξέρχεται καθώς έχει επιτύχει μια σημαντική τιμή p και εκτελούμε τη δοκιμή KPSS στα διαφορετικά δεδομένα και παρατηρούμε τα αποτελέσματα σύμφωνα με τον πίνακα 5.3:

- ADF στατιστικό δοκιμής: -15.5 και p-value: 0
- KPSS στατιστικό δοκιμής: 0.084 και p-value: 0.1

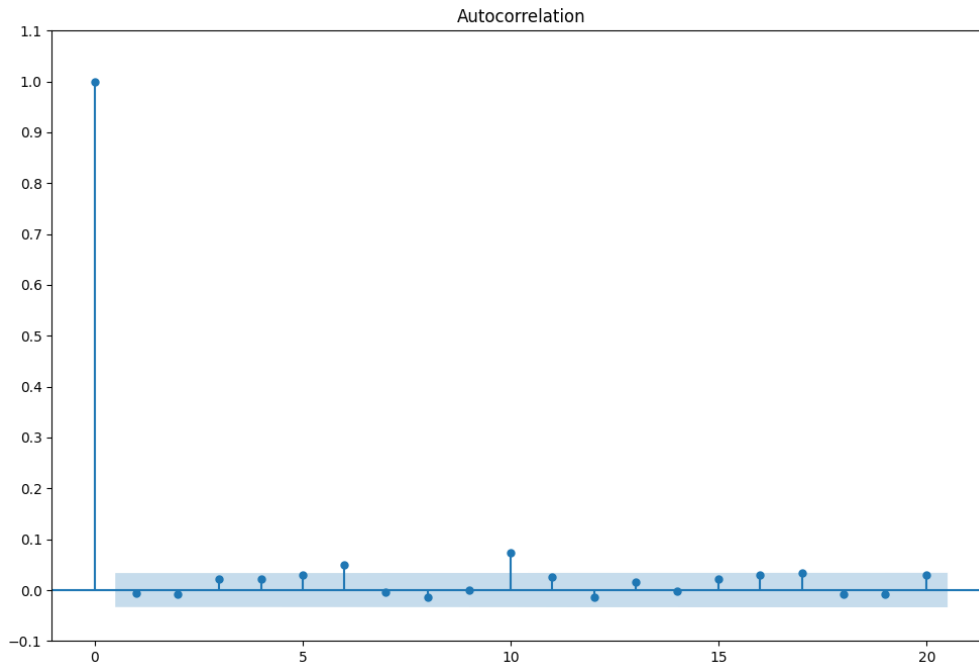
Έχοντας εκτελέσει μια διαφορά πρώτης τάξης στα δεδομένα, η μηδενική υπόθεση έχει απορριφθεί για το ADF, ενώ η μηδενική υπόθεση δεν μπορεί να απορριφθεί για το KPSS, γεγονός που υποδηλώνει ότι η χρονική σειρά είναι πλέον σταθερή. Επιπλέον, αν οπτικοποιήσουμε τη σειρά, τώρα δείχνει σταθερό μέσο όρο και σταθερή διακύμανση (διάγραμμα 5.3).



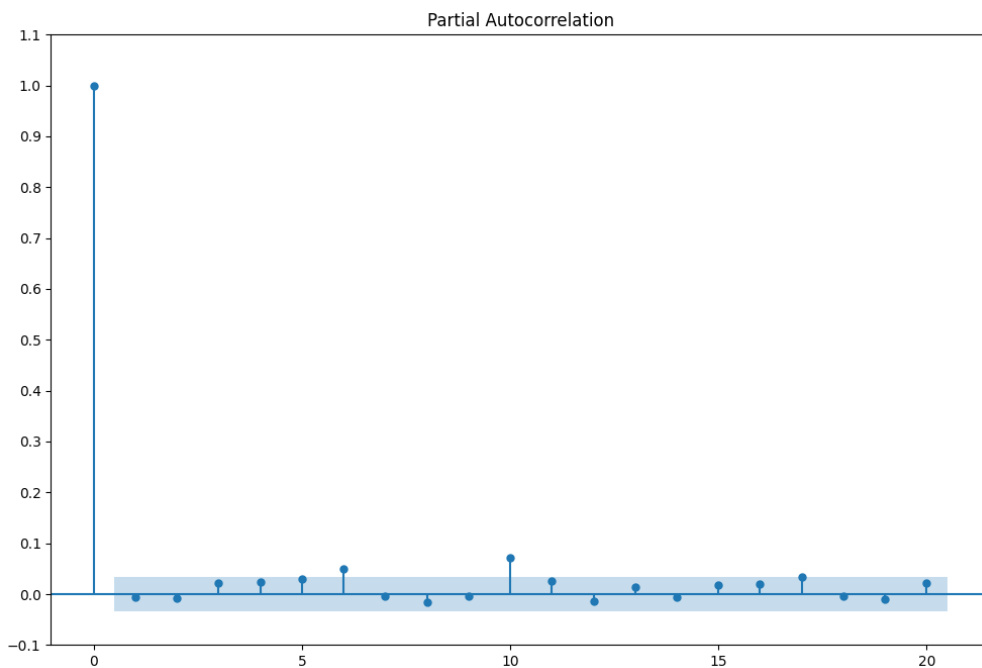
Διάγραμμα 5.3: Οπτικοποίηση δεδομένων χρονοσειράς μετά από διαφοροποίηση για επίτευξη σταθερότητας.

5. ACF & PACF ΈΛΕΓΧΟΣ ΤΥΧΑΙΟΤΗΤΑΣ

Αφού επιτυγχάνεται η σταθερότητα, το επόμενο βήμα είναι να καθοριστούν οι όροι για το τμήμα AR και MA του μοντέλου, με την πραγματοποίηση ελέγχου συσχέτισης των δεδομένων της χρονοσειράς. Ένας τρόπος με τον οποίο μπορούμε να λάβουμε τους όρους είναι να χρησιμοποιήσουμε τις συναρτήσεις γραφικής παράστασης των μοντέλων stats. Το διάγραμμα ACF (5.4) εμφανίζει τους συντελεστές συσχέτισης μεταξύ μιας χρονοσειράς και των τιμών υστέρησής της. Εξηγεί πώς η παρούσα τιμή μιας δεδομένης χρονοσειράς σχετίζεται με προηγούμενες τιμές. Ομοίως, το PACF (διάγραμμα 5.5) εξηγεί τη μερική συσχέτιση μεταξύ της σειράς και των καθυστερήσεων της. Παρατηρώντας και τα δύο διαγράμματα, αναζητούμε σημαντικά σημεία εκτός της σκιασμένης περιοχής (διαστήματος εμπιστοσύνης) καθώς και μια γεωμετρική συσχέτιση/ συμμετρία. Σε περίπτωση που τα δεδομένα δεν είναι τυχαία, το μοντέλο πρόβλεψης ARIMA μπορεί να είναι κατάλληλο.



Διάγραμμα 5.4: Έλεγχος τυχειότητας δεδομένων σε διάγραμμα ACF.



Διάγραμμα 5.5: Έλεγχος τυχειότητας δεδομένων σε διάγραμμα PACF.

Τώρα παρατηρώντας τα διαγράμματα ACF και PACF παρατηρούμε ότι:

- Υπάρχουν ορισμένα σημαντικά σημεία.
- Δεν υπάρχει γεωμετρική ακολουθία

- Υπάρχει μια απότομη πτώση μετά την υστέρηση 0.

Το γράφημα της συνάρτησης αυτοσυσχέτισης «σβήνει» γρήγορα, γεγονός που οδηγεί σε ένδειξη στασιμότητας. Ο καθορισμός των όρων του μοντέλου με τη χρήση αυτών των γραφημάτων είναι περίπλοκος και θα ήταν επιρρεπής σε σφάλματα, καθώς είναι σχετικά πιθανό αυτό που βλέπουμε είναι ένα μοντέλο $ARIMA(0,1,0)$. Η διερεύνηση είναι σημαντικό να γίνει εφόσον μπορεί να ξεγελαστεί κανείς πιστεύοντας ότι αφορά οποιονδήποτε συνδυασμό 6 και 10 για τους όρους AR και MA. Εάν πράγματι το μοντέλο πρόκειται για $ARIMA(0,1,0)$, τα διαφορετικά δεδομένα θα είναι αυτό που είναι γνωστό ως "λευκός θόρυβος" και σημαίνει ότι τα αρχικά δεδομένα είναι αυτό που είναι γνωστό ως "τυχαίος περίπατος".

Εάν οι μεταβλητές είναι ανεξάρτητες και κατανέμονται πανομοιότυπα με μέσο όρο μηδέν, η χρονική σειρά είναι λευκός θόρυβος. Επιπλέον, ένας τυχαίος περίπατος είναι μια σειρά αριθμών όπου κάθε τιμή είναι συνάρτηση της προηγούμενης τιμής. Μια σειρά λευκού θορύβου αφήνεται όταν λαμβάνεται η πρώτη διαφορά από έναν τυχαίο περίπατο.

Καλούμαστε να αντιμετωπίσουμε τα ακόλουθα ζητήματα:

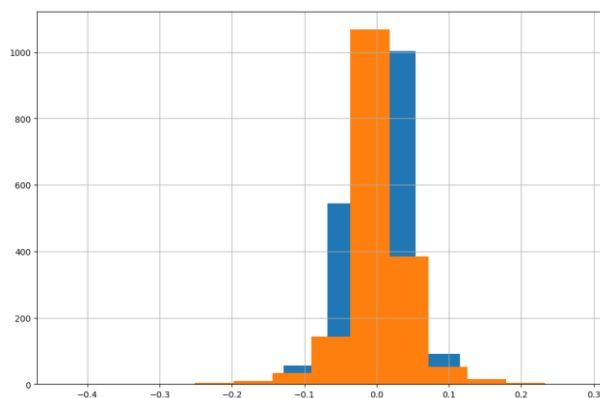
- Πρέπει να ξέρουμε αν η χρονοσειρά μας είναι ένας τυχαίος περίπατος.
- Πρέπει να γνωρίζουμε ποιοι είναι οι όροι μας για το μοντέλο.

Μπορεί ουσιαστικά να βρεθεί λύση στα ζητήματα από κοινού. Με τον διαχωρισμό των δεδομένων στη μέση, παρατηρείται ότι ο μέσος όρος και οι τυπικές αποκλίσεις είναι διαφορετικές όπως και η κατανομή των τιμών είναι ελαφρώς διαφορετική (πίνακας 5.4).

Πίνακας 5.4: Διαχωρισμός δεδομένων στη μέση, εύρεση μέσης τιμές και τυπικών αποκλίσεων κάθε μισού.

```
count    1718.000000
mean      0.002731
std       0.043048
min       -0.251780
25%       -0.011755
50%       0.002391
75%       0.019131
max       0.287099
Name: price, dtype: float64
```

```
count    1718.000000
mean      0.000141
std       0.039223
min       -0.433714
25%       -0.016280
50%       0.001329
75%       0.018126
max       0.176026
Name: price, dtype: float64
```



Διάγραμμα 5.6: Κατανομές τιμών των δύο μισών της χρονοσειράς.

Εκτός από την ερμηνεία των διαγραμμάτων, μπορεί να χρησιμοποιηθεί το τεστ Ljung–Box, το οποίο καθορίζει εάν οποιαδήποτε ομάδα αυτοσυσχετίσεων σε μια χρονοσειρά διαφέρει από το μηδέν. Το τεστ μπορεί να οριστεί ως εξής:

H0: Τα δεδομένα διανέμονται ανεξάρτητα.

H1: Τα δεδομένα δεν κατανέμονται τυχαία. Αντίθετα, παρουσιάζουν σειριακή συσχέτιση.

Πίνακας 5.5: Αποτελέσματα Ljung-Box στα δεδομένα της χρονοσειράς.

	lb_stat	lb_pvalue
20	50.485119	0.000189

Πίνακας 5.6: Αποτελέσματα Ljung-Box στα υπόλοιπα της χρονοσειράς.

	lb_stat	lb_pvalue
20	12.085907	0.913087

Η δοκιμή μπορεί να εφαρμοστεί στα δεδομένα και στα υπολείμματα του προσαρμοσμένου μοντέλου. Εάν η μηδενική υπόθεση απορριφθεί, υπάρχουν στατιστικά στοιχεία ότι τα δεδομένα περιέχουν σειριακή συσχέτιση και δεν είναι λευκός θόρυβος. Από την άλλη πλευρά, όταν εφαρμόζεται η δοκιμή στα υπολείμματα του μοντέλου, ένα καλά προσαρμοσμένο μοντέλο θα πρέπει να εμφανίζει στατιστικά στοιχεία ότι τα υπολείμματα είναι λευκός θόρυβος, καθώς αυτό σημαίνει ότι το μοντέλο έχει καταγράψει όλες τις συσχετισμένες πτυχές στα δεδομένα και όλα τα υπόλοιπα αποτελούν θόρυβο στην χρονοσειρά. Έχοντας εφαρμόσει τη δοκιμή στα δεδομένα, λαμβάνουμε μια τιμή p 0,000189, υπονοώντας ότι έχουμε στατιστικά στοιχεία για την απόρριψη της μηδενικής υπόθεσης και τα δεδομένα δεν κατανέμονται τυχαία και δεν είναι λευκός θόρυβος (πίνακας 5.5).

Στη συνέχεια, εφαρμόζεται η δοκιμή Ljung–Box στα υπολείμματα του μοντέλου και προκύπτει μια τιμή p 0,913087, υπονοώντας ότι δεν γίνεται να απορριφθεί η μηδενική υπόθεση και τα υπολείμματά είναι λευκός θόρυβος (πίνακας 5.6).

6. Εύρεση καλύτερων παραμέτρων μοντέλου

Για την επιλογή του καλύτερου συγκριτικά μοντέλου είναι δυνατή η δοκιμή και η χειροκίνητη προσαρμογή μιας δικής μας ποικιλία μοντέλων. Ωστόσο, μια αυτοματοποιημένη λειτουργία που ονομάζεται "Auto — Arima" προσαρμόζει και δοκιμάζει αυτόματα μια επιλογή μοντέλων και επιστρέφει το μοντέλο με τη χαμηλότερη τιμή AIC. Εφαρμόζοντας τη συνάρτηση, όπως φαίνεται στον πίνακα 5.7, επιστρέφει το αποτέλεσμα με την καλύτερη εφαρμογή να είναι ένα μοντέλο (7,1,8).

Πίνακας 5.7: Χρήση Auto- Arima για σύγκριση και επιλογή του βέλτιστου μοντέλου.

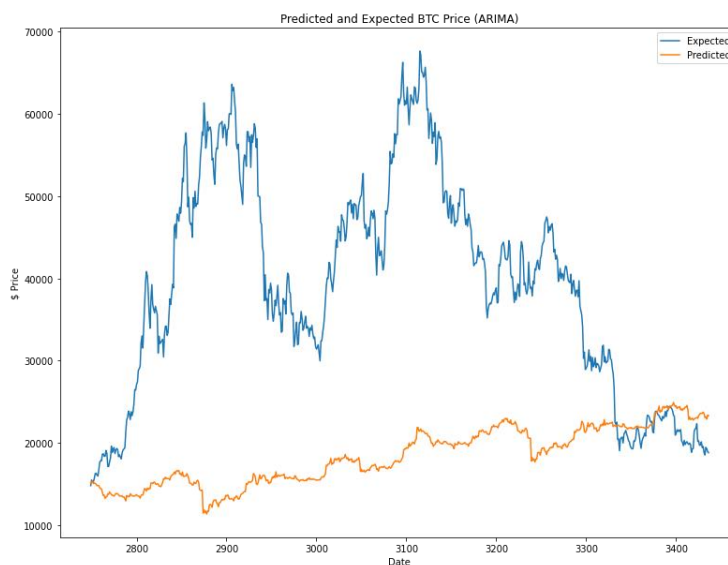
```
Performing stepwise search to minimize aic
ARIMA(10,1,10)(0,0,0)[0] intercept : AIC=-12181.335, Time=12.80 sec
ARIMA(0,1,0)(0,0,0)[0] intercept : AIC=-12163.994, Time=0.19 sec
ARIMA(1,1,0)(0,0,0)[0] intercept : AIC=-12162.100, Time=0.15 sec
ARIMA(0,1,1)(0,0,0)[0] intercept : AIC=-12162.100, Time=0.16 sec
ARIMA(0,1,0)(0,0,0)[0] : AIC=-12161.819, Time=0.22 sec
ARIMA(9,1,10)(0,0,0)[0] intercept : AIC=-12165.009, Time=13.05 sec
ARIMA(10,1,9)(0,0,0)[0] intercept : AIC=-12178.316, Time=13.00 sec
ARIMA(9,1,9)(0,0,0)[0] intercept : AIC=-12185.993, Time=12.20 sec
ARIMA(8,1,9)(0,0,0)[0] intercept : AIC=-12188.604, Time=11.46 sec
ARIMA(7,1,9)(0,0,0)[0] intercept : AIC=-12175.489, Time=10.71 sec
ARIMA(8,1,8)(0,0,0)[0] intercept : AIC=-12192.745, Time=9.31 sec
ARIMA(7,1,8)(0,0,0)[0] intercept : AIC=-12191.366, Time=8.59 sec
ARIMA(8,1,7)(0,0,0)[0] intercept : AIC=-12187.324, Time=8.50 sec
ARIMA(9,1,8)(0,0,0)[0] intercept : AIC=-12190.280, Time=8.32 sec
ARIMA(7,1,7)(0,0,0)[0] intercept : AIC=-12189.739, Time=8.19 sec
ARIMA(9,1,7)(0,0,0)[0] intercept : AIC=-12186.196, Time=10.35 sec
ARIMA(8,1,8)(0,0,0)[0] : AIC=-12193.706, Time=4.90 sec
ARIMA(7,1,8)(0,0,0)[0] : AIC=-12194.041, Time=4.58 sec
ARIMA(6,1,8)(0,0,0)[0] : AIC=-12173.991, Time=4.16 sec
ARIMA(7,1,7)(0,0,0)[0] : AIC=-12193.610, Time=4.44 sec
ARIMA(7,1,9)(0,0,0)[0] : AIC=-12178.184, Time=5.64 sec
ARIMA(6,1,7)(0,0,0)[0] : AIC=-12179.553, Time=3.93 sec
ARIMA(6,1,9)(0,0,0)[0] : AIC=-12168.279, Time=5.31 sec
ARIMA(8,1,7)(0,0,0)[0] : AIC=-12190.071, Time=4.72 sec
ARIMA(8,1,9)(0,0,0)[0] : AIC=-12193.185, Time=6.24 sec

Best model: ARIMA(7,1,8)(0,0,0)[0]
Total fit time: 171.116 seconds
```

7. Εφαρμογή μοντέλου

Εφαρμόζουμε και δοκιμάζουμε το μοντέλο, χρησιμοποιώντας σαν δοκιμή το 20% της χρονοσειράς και σαν εκπαίδευση το πρώτο 80%. Παρατηρείται ότι ακόμα και το καλύτερα

εφαπτόμενο μοντέλο ARIMA στην χρονοσειρά, δεν καλύπτει επάξια τις μελλοντικές τιμές δοκιμής.



Διάγραμμα 5.7: Τεστ του μοντέλου Πρόβλεψης της τιμής του BTC- μέθοδος ARIMA.

Όπως δείχνει το διάγραμμα 5.7, το μοντέλο φαίνεται να μην είναι καλό στην πρόβλεψη της τάσης της σειράς, αλλά ούτε και στις τιμές της πρόβλεψης, οι οποίες είναι σταθερά κάτω από την πραγματική τιμή. Η υποπρόβλεψη φαίνεται να συμβαίνει περισσότερο κατά τη διάρκεια των τεράστιων πτητικών αιχμών προς τα πάνω. Για την αξιολόγηση του μοντέλου και την μελλοντική σύγκρισή του με άλλα απαιτείται να υπολογιστεί ο δείκτης RMSE. Το ριζικό μέσο τετραγωνικό σφάλμα (RMSE) έχει χρησιμοποιηθεί ως τυπική στατιστική μέτρηση για τη μέτρηση της απόδοσης του μοντέλου και έχει αποδειχτεί επιστημονικά ως η πιο κατάλληλη μέθοδος (Chai & Draxler, 2014).

Με την βοήθεια των βιβλιοθηκών της rpython βρέθηκε ότι ο δείκτης RMSE του μοντέλου είναι ίσος: Test RMSE: 25369,142. Όπως γίνεται αντιληπτό και διαγραμματικά, πρόκειται για μεγάλη τιμή σφάλματος. Η τιμή του δείκτη RMSE δεν εξασφαλίζει ότι το μοντέλο δεν είναι ικανό να εκτιμήσει σωστά τις μελλοντικές προβλέψεις. Πριν απορριφθεί τελείως η μέθοδος ARIMA για την συγκεκριμένη εκτίμηση, θα πρέπει να διερευνηθεί εκ νέου η χρονοσειρά.

8. Επανεξέταση μοντέλου

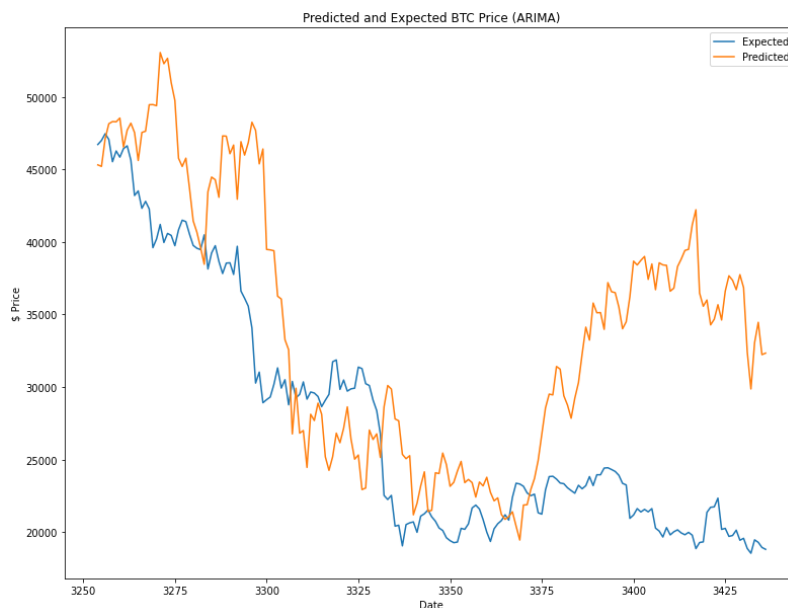
Με μία πιο προσεκτική ματιά πίσω σε όλες τις διακυμάνσεις των τιμών με την πάροδο του χρόνου, είναι αξιοσημείωτο ότι πριν από το 2017, οι κινήσεις των τιμών δεν ήταν τόσο δυναμικές όσο από το 2017 και μετά. Επομένως τα παλιότερα δεδομένα μπορεί να μην είναι

ακριβής αναπαράσταση του τρόπου συμπεριφοράς του Bitcoin επί του παρόντος. Ίσως η κατάργηση των προηγούμενων ετών και η εξέταση του 2017 και μετά θα είναι πιο ενδεικτική και θα επιβεβαιώσει την υποψία ότι αυτή η χρονοσειρά είναι ένας τυχαίος περίπατος. Έτσι, αυτό που μπορούμε να κάνουμε είναι να προχωρήσουμε και να κάνουμε την προηγούμενη λογική στο νέο μειωμένο σύνολο δεδομένων.

9. Δημιουργία νέου μοντέλου και αξιολόγηση

Με την επανάληψη τις διαδικασίας εκ νέου αυτή την φορά για τις τιμές από το 2017 και μετά, εξετάζουμε και πάλι το μοντέλο μας. Περίοδος δοκιμής και εκπαίδευσης του μοντέλου θεωρείται το 90% της χρονοσειράς, ενώ δοκιμής της πρόβλεψης 10%. Αξιολογώντας σταθερότητα (με επιτυχία εφαρμογής μετά από μία διαφοροποίηση) και μη ύπαρξης τυχαιότητας, εφαρμόζεται και πάλι το Auto – Arima.

Με την βοήθεια της Python, το καλύτερο δυνατό μοντέλο βρέθηκε να είναι το: $ARIMA(2,1,0)(0,0,0)[0]$.



Διάγραμμα 5.8: Τεστ του βελτιωμένου μοντέλου Πρόβλεψης της τιμής του BTC- μέθοδος ARIMA.

Όπως δείχνει το διάγραμμα 5.8, το μοντέλο φαίνεται να είναι καλό στην πρόβλεψη της τάσης της σειράς. Η υποπρόβλεψη φαίνεται να συμβαίνει περισσότερο κατά τη διάρκεια των τεράστιων πτητικών αιχμών προς τα πάνω. Οι καθοδικές κλίσεις φαίνεται να έχουν πιο ακριβείς προβλέψεις από τις ανοδικές τάσεις. Αν και το μοντέλο δεν φαίνεται έμπειρο στην πρόβλεψη

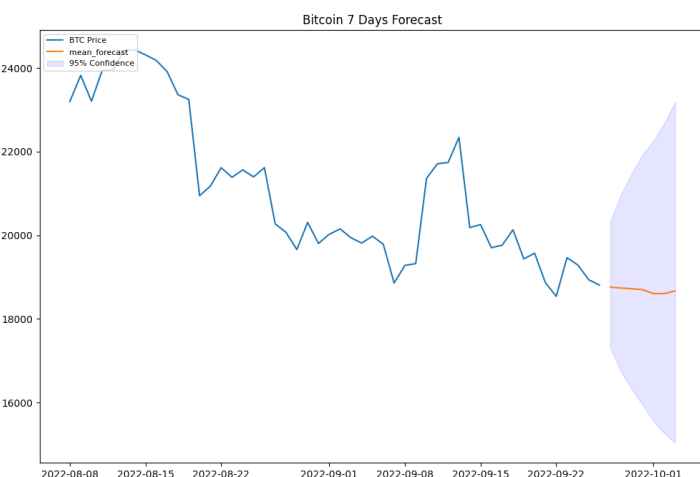
της ίδιας της τιμής, η ικανότητά του να προβλέπει τάσεις όπως αυτή δείχνει ότι μπορεί να υπάρχει πιθανότητα πρόβλεψης τιμών κρυπτογράφησης κάποια στιγμή στο μέλλον.

Η αξιολόγησή του μέσω της τιμής του δείκτη RMSE δίνει: Test RMSE: 9625,666.

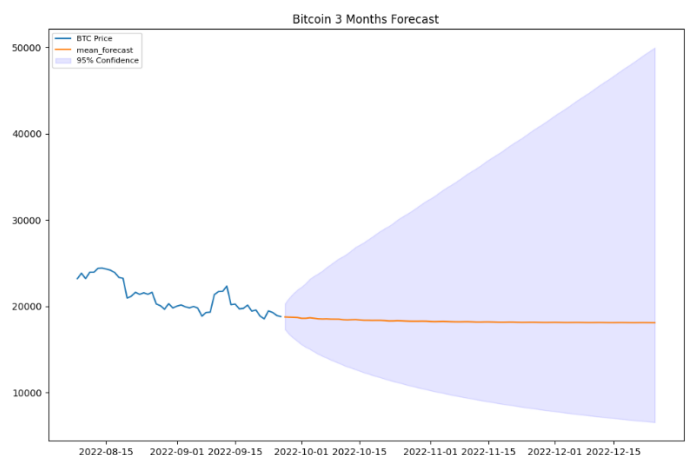
10. Προβλέψεις μελλοντικών τιμών

Εφόσον βρέθηκε το κατάλληλο από τα μοντέλα ARIMA, μπορούν να πραγματοποιηθούν οι μελλοντικές προβλέψεις. Στην συνέχεια προβλέπεται η τιμή του BTC για την επόμενη εβδομάδα (27/09/2022 – 4/10/2022) και οπτικοποιείται το αποτέλεσμα.

Το παρακάτω διάγραμμα 5.9 δείχνει την πρόβλεψη του μοντέλου μας για τις επόμενες επτά ημέρες και απεικονίζει το εύρος τιμών που μπορούμε να περιμένουμε με 95% εμπιστοσύνη. Η πορτοκαλί γραμμή είναι η μέση πρόβλεψη, ενώ η σκιασμένη μωβ περιοχή έχει άνω και κάτω διαστήματα εμπιστοσύνης. Ενδεικτικά παρουσιάζεται και το διάγραμμα 5.10, όπου φαίνεται η πρόβλεψη σε έναν πιο μακροπρόθεσμο ορίζοντα τριών μηνών.



Διάγραμμα 5.9: Πρόβλεψη της τιμής BTC για μία εβδομάδα (27/09/2022 - 4/10/2022) με διάστημα εμπιστοσύνης 95%.



Διάγραμμα 5.10: Πρόβλεψη της τιμής BTC για τρεις μήνες (27/09/2022 - 27/12/2022) με διάστημα εμπιστοσύνης 95%.

5.2.2 Μέθοδος LSTM

1) Επεξεργασία δεδομένων

Για την χρήση μεθόδου LSTM απαιτείται σωστή προεπεξεργασία των δεδομένων. Πρώτα, πρέπει να δημιουργηθεί μια «ακολουθία». Τα LSTM χρησιμοποιούν σειρές ή «διαδοχικά» δεδομένα (σε αυτήν την περίπτωση, η ακολουθία είναι χρόνος, πιο συγκεκριμένα «ημέρες»). Εδώ, η ακολουθία ορίζεται σε 100, που σημαίνει ότι το μοντέλο θα προσπαθήσει να μάθει

μακροπρόθεσμες εξαρτήσεις μεταξύ χρονικών βημάτων των 100 ημερών κάθε φορά. Αυτό μπορεί να αλλάξει ανάλογα με τον πειραματισμό.

2) Διαχωρισμός

Στη συνέχεια, πρέπει να γίνει διαχωρισμός των δεδομένων σε εκπαίδευση και δοκιμή. Μετά την αναμόρφωση των δεδομένων σε πίνακες 3 παραγόντων που μπορούν να τροφοδοτηθούν στο μοντέλο, τα δεδομένα χωρίζονται σε 80% εκπαίδευση και 20% δοκιμή. Αυτό αφήνει 2.669 περιπτώσεις για εκπαίδευση και 668 για δοκιμή.

3) Δημιουργία μοντέλου

Για την δημιουργία του μοντέλου χρειάζεται εξ' αρχής ένα βασικό πρότυπο με κυρίως προεπιλεγμένες παραμέτρους. Αυτό μπορεί να επεκταθεί εύκολα ή να προσαρμοστεί με πειράματα. Τα μοντέλα νευρωνικών δικτύων είναι ενσωματωμένα διαδοχικά επίπεδα, με διαθέσιμες διαφορετικές παραμέτρους και στρατηγικές βελτιστοποίησης.

Αυτό το μοντέλο θα έχει επίπεδα LSTM, με προεπιλογή 32 κρυφά επίπεδα. Έχει επίσης οριστεί ποσοστό εγκατάλειψης 0,2. Το Dropout είναι μια τεχνική όπου οι τυχαία επιλεγμένοι νευρώνες αγνοούνται κατά τη διάρκεια της προπόνησης, αυτή είναι μια τεχνική τακτοποίησης που έχει σχεδιαστεί για να σταματήσει την υπερβολική προσαρμογή). Ένα ποσοστό εγκατάλειψης 0,2 σημαίνει ότι το 20% των νευρώνων αγνοείται, επομένως το μοντέλο δεν μαθαίνει τα δεδομένα στο σημείο που ταιριάζουν ακριβώς πάνω τους, καθιστώντας τις προβλέψεις μας άχρηστες. Το τελικό στρώμα εξόδου είναι ένα πυκνό στρώμα. Αυτό συγχωνεύει όλα τα επίπεδα μας σε ένα μικρότερο σύνολο επιλογών, παρέχοντας έτσι το προβλεπόμενο αποτέλεσμα.

Στη συνέχεια, το μοντέλο συντάσσεται, η απώλεια μοντέλου μετράται σε MSE, το πρότυπο για γραμμικά μοντέλα για προβλέψεις. Επιλέχθηκε ο βελτιστοποιητής ADAM (μια συνάρτηση που αλλάζει τα βάρη ανάλογα για να βελτιώσει την ακρίβεια), αλλά υπάρχουν πολλά πιθανά βελτιστοποιητικά προς χρήση, ανάλογα με το περιβάλλον.

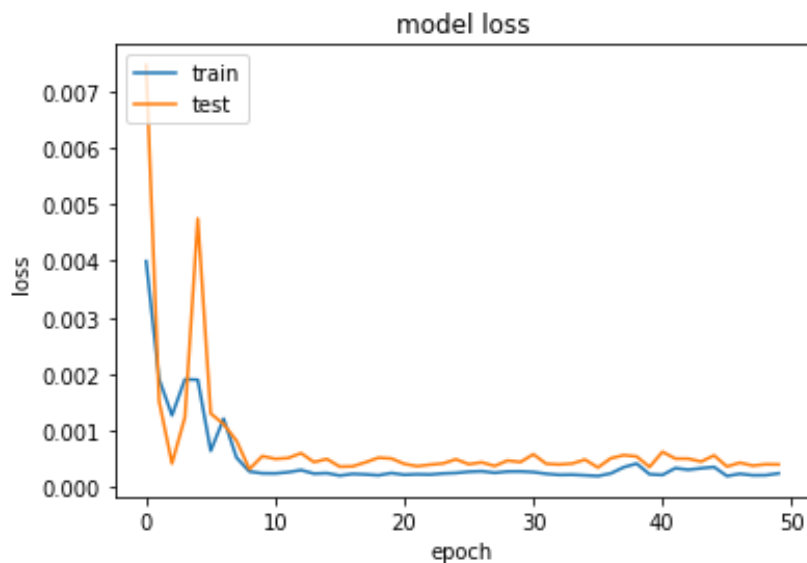
Πίνακας 5.8: Βασικό μοντέλο LSTM – πρόβλεψη BTC.

```
Model: "sequential_12"  
-----  
Layer (type)                Output Shape                Param #  
-----  
lstm_36 (LSTM)              (None, 99, 32)            4352  
lstm_37 (LSTM)              (None, 99, 32)            8320  
lstm_38 (LSTM)              (None, 32)                 8320  
dense_12 (Dense)           (None, 1)                  33  
-----  
Total params: 21,025  
Trainable params: 21,025  
Non-trainable params: 0
```

Τώρα το μοντέλο είναι εκπαιδευμένο. Ένα μέγεθος παρτίδας 300 σημαίνει ότι είναι ο αριθμός των περιπτώσεων που τροφοδοτούνται μέσω του μοντέλου ανά πάσα στιγμή. Οι εποχές ορίζονται σε 50. Διαγράφεται επίσης ένα σύνολο επικύρωσης 10%, το οποίο θα μας επιτρέψει να επικυρώνουμε με μεγαλύτερη ακρίβεια τα αποτελέσματα.

4) Προβολή μοντέλου

Η συνάρτηση απώλειας σε ένα νευρωνικό δίκτυο δείχνει τη διαφορά μεταξύ του αναμενόμενου αποτελέσματος και του αποτελέσματος που παράγεται από το μοντέλο. Στο παρακάτω διάγραμμα 5.11 παρατηρείται η απόδοση του μοντέλου στην «προπόνηση».

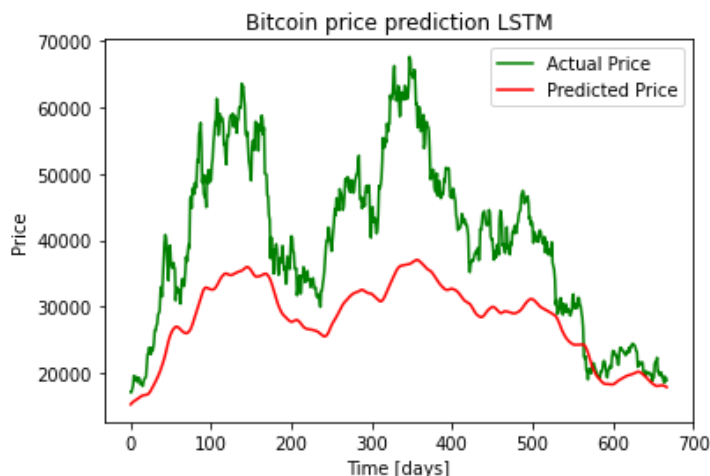


Διάγραμμα 5.11: Εφαρμογή του επιλεγμένου μοντέλου LSTM στα δεδομένα της χρονοσειράς.

Η εφαρμογή της γραμμής απώλειας της δοκιμής φαίνεται να εφάπτεται και να ακολουθεί την τάση της γραμμής δοκιμής. Αυτό είναι ένα σημάδι επαρκούς προσαρμογής.

5) Δοκιμή του μοντέλου

Όπως δείχνει το διάγραμμα 5.12, το μοντέλο φαίνεται να είναι αρκετά καλό στην πρόβλεψη της τάσης της σειράς, αλλά είναι σταθερά κάτω από την πρόβλεψη. Η υποπρόβλεψη φαίνεται να συμβαίνει περισσότερο κατά τη διάρκεια των τεράστιων πτητικών αιχμών προς τα πάνω. Το μοντέλο φαίνεται να ενεργεί συντηρητικά, υπό κυρίως της προβλέψεως. Οι καθοδικές κλίσεις φαίνεται να έχουν πιο ακριβείς προβλέψεις από τις ανοδικές τάσεις. Αν και το μοντέλο δεν φαίνεται έμπειρο στην πρόβλεψη της ίδιας της τιμής, η ικανότητά του να προβλέπει τάσεις όπως αυτή δείχνει ότι μπορεί να υπάρχει πιθανότητα πρόβλεψης τιμών κρυπτογράφησης κάποια στιγμή στο μέλλον.



Διάγραμμα 5.12: Τεστ του μοντέλου Πρόβλεψης της τιμής του BTC- μέθοδος LSTM.

6) Αξιολόγηση μοντέλου

Τέλος, και πάλι ο υπολογισμός της τιμής RMSE θα βοηθήσει στην αξιολόγηση και τελική σύγκριση του μοντέλου LSTM με τα υπόλοιπα. Με την βοήθεια της βιβλιοθήκης “math” στην Python βρέθηκε η τιμή RMSE: Train data RMSE: 866,30.

5.2.3 Μέθοδος Random Forest

1. Διαχωρισμός δεδομένων

Πρώτο στάδιο έναρξης της διαδικασίας πρόβλεψης είναι να διαχωριστούν τα δεδομένα που θα δοκιμαστούν για πρόβλεψη. Το μοντέλο θα εκπαιδευτεί κατά 90% στα δεδομένα της χρονοσειράς και θα δοκιμαστεί στο επόμενο 10% για προβλέψεις (διάγραμμα 5.13). Αναλυτικότερα πρόκειται για:

- Μέρες προπόνησης (Train dates) : 28/04/2013 έως 17/10/2021 (n=3095 μέρες)
- Μέρες δοκιμής πρόβλεψης (Test dates) : 18/10/2021 έως 26/09/2022 (n=344 μέρες)



Διάγραμμα 5.13: Μέρες εκπαίδευσης και δοκιμής του μοντέλου Random Forest για πρόβλεψη BTC.

2. Δημιουργία δοκιμαστικού μοντέλου

Με την κλάση `ForecasterAutoreg`, δημιουργείται ένα μοντέλο και εκπαιδεύεται από έναν αναδρομητή `RandomForestRegressor` με χρονικό παράθυρο 364 καθυστερήσεων. Αυτό σημαίνει ότι το μοντέλο χρησιμοποιεί τις προηγούμενες 364 ημέρες ως βάση για την πρόβλεψη.



Διάγραμμα 5.14: Δοκιμή προβλέψεων τιμής BTC με το μοντέλο Random Forest.

Για να εξεταστεί το μοντέλο και να συγκριθεί με τα υπόλοιπα, απαιτείται και πάλι ο υπολογισμός του δείκτη RMSE.

Test RMSE: 23.838.

3. Εύρεση καλύτερου μοντέλου Random Forest

Το εκπαιδευμένο ForecasterAutoreg χρησιμοποιεί ένα παράθυρο χρόνου 6 υστερήσεων και ένα μοντέλο Random Forest με τις προεπιλεγμένες υπερπαραμέτρους. Ωστόσο, οι προεπιλεγμένες τιμές δεν είναι οι κατάλληλες. Προκειμένου να εντοπιστεί ο καλύτερος συνδυασμός καθυστερήσεων και υπερπαραμέτρων, η βιβλιοθήκη Skforecast παρέχει τη συνάρτηση `grid_search_forecaster`. Συγκρίνει τα αποτελέσματα που λαμβάνονται με κάθε διαμόρφωση μοντέλου.

Number of models compared: 12

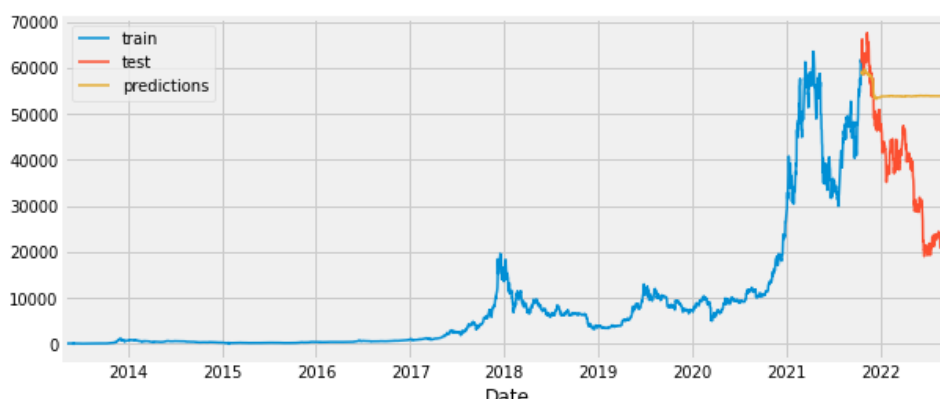
```
loop lags_grid: 0%| | 0/2 [00:00<?, ?it/s]
loop param_grid: 0%| | 0/6 [00:00<?, ?it/s]
loop param_grid: 17%|██████████| 1/6 [00:12<01:01, 12.31s/it]
loop param_grid: 33%|██████████| 2/6 [01:50<04:10, 62.66s/it]
loop param_grid: 50%|██████████| 3/6 [02:01<01:57, 39.03s/it]
loop param_grid: 67%|██████████| 4/6 [03:47<02:11, 65.75s/it]
loop param_grid: 83%|██████████| 5/6 [03:59<00:46, 46.41s/it]
loop param_grid: 100%|██████████| 6/6 [05:59<00:00, 71.11s/it]
loop lags_grid: 50%|██████████| 1/2 [05:59<05:59, 359.04s/it]
loop param_grid: 0%| | 0/6 [00:00<?, ?it/s]
loop param_grid: 17%|██████████| 1/6 [01:13<06:08, 73.69s/it]
loop param_grid: 33%|██████████| 2/6 [13:18<30:27, 456.80s/it]
loop param_grid: 50%|██████████| 3/6 [15:03<14:48, 296.04s/it]
loop param_grid: 67%|██████████| 4/6 [32:37<19:51, 595.51s/it]
loop param_grid: 83%|██████████| 5/6 [35:31<07:23, 443.33s/it]
loop param_grid: 100%|██████████| 6/6 [1:03:56<00:00, 872.22s/it]
loop lags_grid: 100%|██████████| 2/2 [1:09:55<00:00, 2097.67s/it]
```

Εικόνα 5.2: Χρήση της συνάρτησης `grid search forecaster` για την εύρεση καλύτερων υπερπαραμέτρων.

Με την βοήθεια του `Skforecast` (εικόνα 5.2) βρέθηκαν ο καλύτερος συνδυασμός παραμέτρων για το συγκεκριμένο μοντέλο: `{'max_depth': 3, 'n_estimators': 1000}`

Δοκιμή μοντέλου και αξιολόγηση

Η δοκιμή του μοντέλου `Random Forest(3,1000)` γίνεται αρχικά διαγραμματικά. Παρατηρούμε στο διάγραμμα 5.15 ότι το μοντέλο προβλέπει τις τιμές με ακρίβεια μόνο για ένα μικρό χρονικό διάστημα. Όταν ο ορίζοντας πρόβλεψης μεγαλώνει, οι εκτιμήσεις του συγκεκριμένου μοντέλου αστοχούν.



Διάγραμμα 5.15: Εύρεση βέλτιστου μοντέλου προβλέψεων τιμής BTC- `Random Forest`.

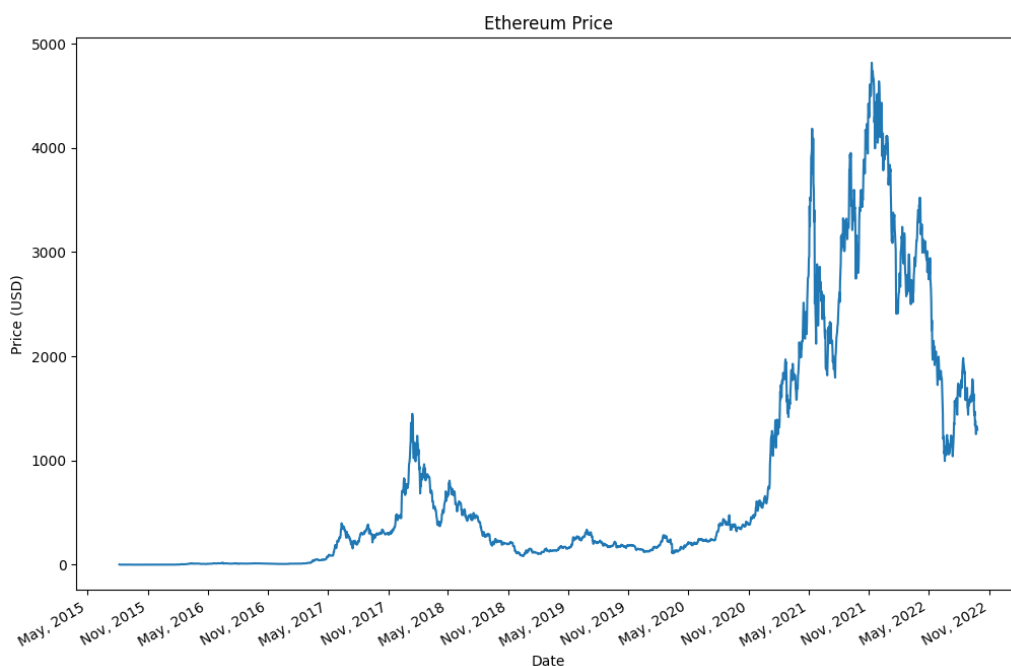
Για να εξεταστεί το μοντέλο απαιτείται και πάλι ο υπολογισμός του δείκτη `RMSE`.

Test `RMSE`: 21.279,875.

5.3 Πρόβλεψη τιμής Ethereum

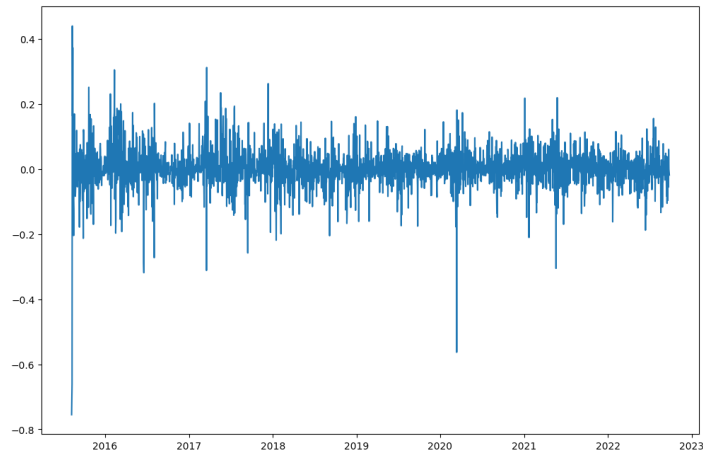
5.3.1 Μέθοδος ARIMA

Η χρονοσειρά της τιμής του Ethereum (διάγραμμα 5.16) ακολουθεί παρόμοια πορεία τάσης με αυτή του BTC. Από τον Ιούνιο του 2017 παρατηρείται απότομη αύξηση της τιμής του νομίσματος. Αφού η πρώτη ανατάραξη των τιμών έκανε τον κύκλο της, μία νέα πορεία αύξησης τιμών έρχεται στο τέλος του 2020 και κορυφώνεται τον Νοέμβριο του 2021. Από τότε έως σήμερα το νόμισμα ακολουθεί πτωτική πορεία όμως και πάλι η τιμή του είναι παραπάνω από αξιοσημείωτη συγκριτικά με την πορεία του.



Διάγραμμα 5.16: : Αναπαράσταση ιστορικών στοιχείων χρονοσειράς για την τιμή του Ethereum.

Η έντονη τάση της χρονοσειράς καθιστά σαφές την ανάγκη για διαφοροποιήσεις προκειμένου να μετατρέψουμε την ακολουθία σε στάσιμη. Μετά από μία διαφοροποίηση, η σταθερότητα επέρχεται όπως φαίνεται και στο διάγραμμα 5.17.



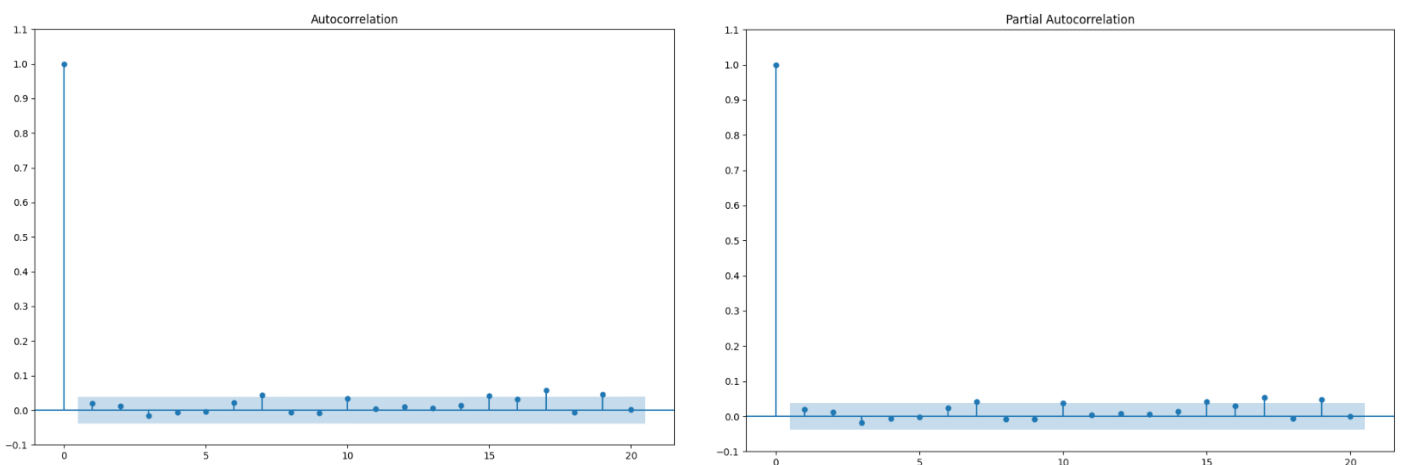
Διάγραμμα 5.17: Οπτικοποίηση δεδομένων χονοσειράς μετά από διαφοροποίηση για επίτευξη σταθερότητας.

Μετά από μία επανάληψη για διαφοροποίηση των δεδομένων, ο βρόχος εξέρχεται καθώς έχει επιτύχει μια σημαντική τιμή p και εκτελούμε τη δοκιμή KPSS στα διαφορετικά δεδομένα και παρατηρούμε τα αποτελέσματα:

- ADF στατιστικό δοκιμής: -25.33 και p -value: 0
- KPSS στατιστικό δοκιμής: 0.1044 και p -value: 0.1

Έχοντας εκτελέσει μια διαφορά πρώτης τάξης στα δεδομένα, η μηδενική υπόθεση έχει απορριφθεί για το ADF, ενώ η μηδενική υπόθεση δεν μπορεί να απορριφθεί για το KPSS, γεγονός που υποδηλώνει ότι η χρονική σειρά είναι πλέον σταθερή. Επιπλέον, αν οπτικοποιήσουμε τη σειρά, τώρα δείχνει σταθερό μέσο όρο και σταθερή διακύμανση.

Ακολουθεί ο έλεγχος τυχαιότητας με την μέθοδο ACF και PACF (διάγραμμα 5.18).



Διάγραμμα 5.18: Έλεγχος τυχαιότητας δεδομένων σε διάγραμμα ACF και PACF.

Παρατηρώντας τα διαγράμματα ACF και PACF παρατηρούμε ότι:

- Υπάρχουν ορισμένα σημαντικά σημεία.
- Δεν υπάρχει γεωμετρική ακολουθία
- Υπάρχει μια απότομη πτώση μετά την υστέρηση 0.

Στη συνέχεια, εφαρμόζουμε τη δοκιμή Ljung–Box για να προσδιορίσουμε τυχόν τυχαιότητα στα δεδομένα και στα υπολείμματα του μοντέλου. Για τα δεδομένα της χρονοσειράς παίρνουμε μια τιμή p-value 0,001572, υπονοώντας ότι υπάρχουν στατιστικά στοιχεία για την απόρριψη της μηδενικής υπόθεσης άρα τα δεδομένα δεν κατανέμονται τυχαία και δεν είναι λευκός θόρυβος.

Αφού ολοκληρώνεται ο έλεγχος σταθερότητας και τυχαιότητας, γίνεται με την βοήθεια του Auto – Arima η επιλογή του καλύτερου μοντέλου για την πρόβλεψη της συγκεκριμένης χρονοσειράς. Το καλύτερο μοντέλο σύμφωνα με το πρόγραμμά μας είναι το ARIMA(0,1,0)(0,0,0). Το πρόβλημα με το συγκεκριμένο μοντέλο είναι ότι αντιστοιχεί σε μοντέλο τυχαίου περιπάτου.

Εάν η σειρά Y δεν είναι ακίνητη, το απλούστερο δυνατό μοντέλο για αυτήν είναι ένα μοντέλο τυχαίας βόλτας, το οποίο μπορεί να θεωρηθεί ως περιοριστική περίπτωση ενός μοντέλου AR(1) στο οποίο ο αυτοπαλινδρομικός συντελεστής είναι ίσος με 1, δηλαδή μια σειρά με απείρως αργή μέση αναστροφή. Η εξίσωση πρόβλεψης για αυτό το μοντέλο μπορεί να γραφτεί ως εξής:

Εξίσωση 5: Εξίσωση πρόβλεψης για μοντέλο ARIMA(0,1,0).

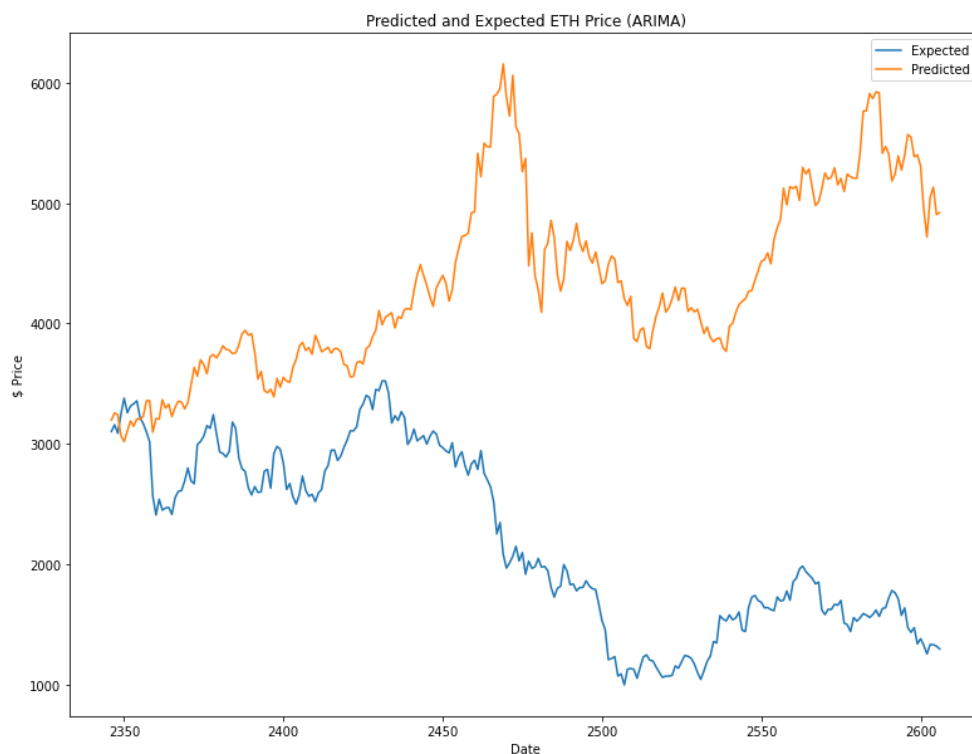
$$\hat{Y}_t = \mu + Y_{t-1}$$

...όπου ο σταθερός όρος είναι η μέση μεταβολή από περίοδο σε περίοδο (δηλαδή η μακροπρόθεσμη μετατόπιση) στο Y . Αυτό το μοντέλο θα μπορούσε να προσαρμοστεί ως μοντέλο παλινδρόμησης χωρίς τομή στο οποίο η πρώτη διαφορά του Y είναι η εξαρτημένη μεταβλητή. Δεδομένου ότι περιλαμβάνει (μόνο) μια μη εποχική διαφορά και έναν σταθερό όρο, ταξινομείται ως "μοντέλο ARIMA(0,1,0) με σταθερά." Το μοντέλο τυχαίου περιπάτου χωρίς μετατόπιση θα ήταν ένα μοντέλο ARIMA(0,1,0) χωρίς σταθερά.

Με μία πιο προσεκτική ματιά πίσω σε όλες τις διακυμάνσεις των τιμών με την πάροδο του χρόνου, είναι αξιοσημείωτο ότι πριν από τον Ιούλιο του 2017, οι κινήσεις των τιμών δεν ήταν

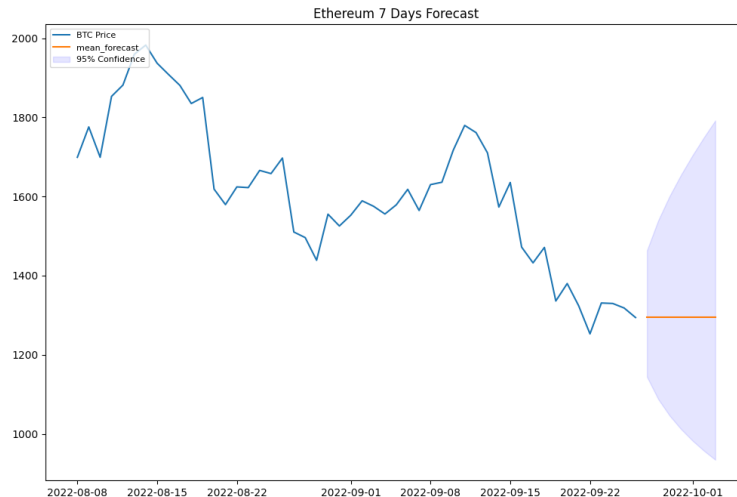
τόσο δυναμικές όσο από το 2017 και μετά. Επομένως τα παλιότερα δεδομένα μπορεί να μην είναι ακριβής αναπαράσταση του τρόπου συμπεριφοράς του Ethereum επί του παρόντος. Ίσως η κατάργηση των προηγούμενων ετών και η εξέταση του 2017 και μετά θα είναι πιο ενδεικτική και θα επιβεβαιώσει την υποψία ότι αυτή η χρονοσειρά είναι ένας τυχαίος πείπατος. Έτσι, αυτό που μπορούμε να κάνουμε είναι να προχωρήσουμε και να κάνουμε την προηγούμενη λογική στο νέο μειωμένο σύνολο δεδομένων.

Με την βοήθεια και πάλι του AUTO-ARIMA και το νέο μειωμένο σύνολο δεδομένων χρονοσειράς καταλήγουμε ότι το μοντέλο ARIMA(2,1,0) είναι το καλύτερο δυνατό. Με δοκιμή του μοντέλου σε ποσοστό 90% δεδομένα χρονοσειράς και 10% εκπαίδευσης προκύπτει το ακόλουθο διάγραμμα 5.19.



Διάγραμμα 5.19: Τεστ του μοντέλου Πρόβλεψης της τιμής του ETH- μέθοδος ARIMA.

Το RMSE στο παρόν μοντέλο προκύπτει ίσο με 2.418. Η τιμή είναι ιδιαίτερος υψηλή κάτι που δικαιολογείται και οπτικά στο διάγραμμα. Παρατηρείται ότι η δοκιμή των προβλέψεων χαρακτηρίζεται από έντονη υποπρόβλεψη.



Διάγραμμα 5.20: Πρόβλεψη της τιμής Ethereum για μία εβδομάδα (27/09/2022 - 4/10/2022) με διάστημα εμπιστοσύνης 95%

Το παραπάνω διάγραμμα 5.20 δείχνει την πρόβλεψη του μοντέλου μας για τις επόμενες επτά ημέρες και απεικονίζει το εύρος τιμών που μπορούμε να περιμένουμε με 95% εμπιστοσύνη. Η πορτοκαλί γραμμή είναι η μέση πρόβλεψη, ενώ η σκιασμένη μωβ περιοχή έχει άνω και κάτω διαστήματα εμπιστοσύνης.

5.3.2 Μέθοδος LSTM

Η ακολουθία ορίζεται σε 100, που σημαίνει ότι το μοντέλο θα προσπαθήσει να μάθει μακροπρόθεσμες εξαρτήσεις μεταξύ χρονικών βημάτων των 100 ημερών κάθε φορά. Αυτό μπορεί να αλλάξει ανάλογα με τον πειραματισμό.

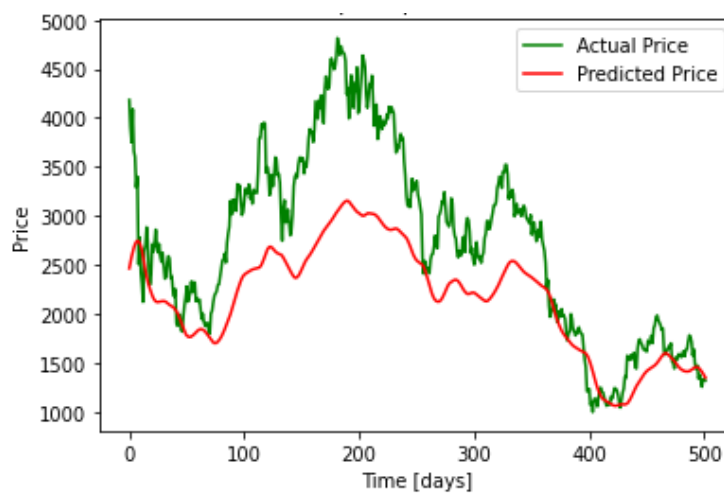
Μετά την αναμόρφωση των δεδομένων σε πίνακες 3 παραγόντων που μπορούν να τροφοδοτηθούν στο μοντέλο, τα δεδομένα χωρίζονται σε 80% εκπαίδευση και 20% δοκιμή. Αυτό αφήνει 2.005 περιπτώσεις για εκπαίδευση και 502 για δοκιμή.

Αυτό το μοντέλο θα έχει επίπεδα LSTM, με προεπιλογή 32 κρυφά επίπεδα. Έχει επίσης οριστεί ποσοστό εγκατάλειψης 0,2. Επιπλέον, το ποσοστό εγκατάλειψης ορίζεται σε 20%.

Στη συνέχεια, το μοντέλο συντάσσεται, η απώλεια μοντέλου μετράται σε MSE, το πρότυπο για γραμμικά μοντέλα για προβλέψεις. Επιλέχθηκε ο βελτιστοποιητής ADAM.

Πίνακας 5.9: Βασικό μοντέλο LSTM - πρόβλεψη ETH.

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 99, 32)	4352
lstm_1 (LSTM)	(None, 99, 32)	8320
lstm_2 (LSTM)	(None, 32)	8320
dense (Dense)	(None, 1)	33



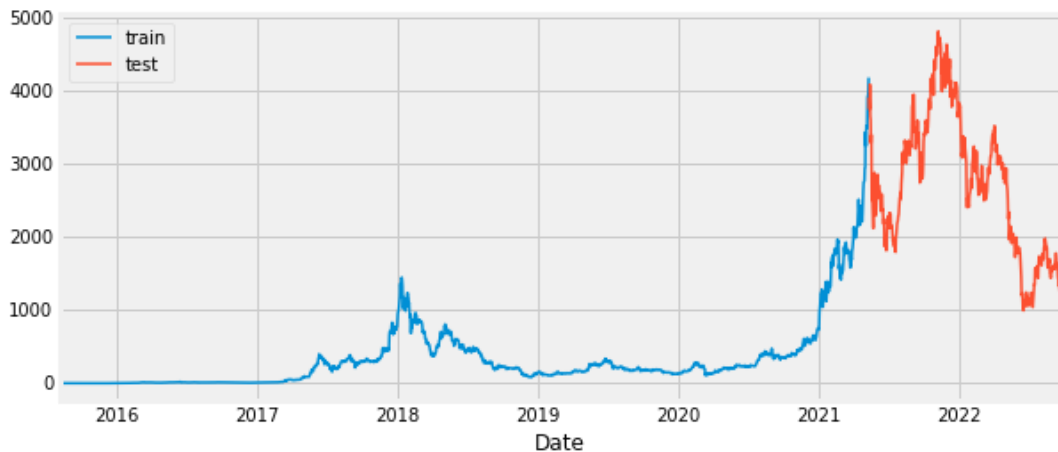
Διάγραμμα 5.21: Τεστ του μοντέλου Πρόβλεψης της τιμής του ETH- μέθοδος LSTM.

Όπως παρατηρείται στο διάγραμμα 5.21, το μοντέλο εφαρμόζει καλά στην χρονοσειρά. Οι ραγδαίες αυξήσεις τις τιμές και πάλι δεν εντοπίζονται με ακρίβεια όπως τις πτωτικές τιμές, όμως μπορούμε να πάρουμε ασφαλή συμπεράσματα για την τάση της σειράς. Το RMSE προκύπτει ίσο με 143,04.

5.3.3 Μέθοδος Random Forest

Το μοντέλο θα εκπαιδευτεί κατά 80% στα δεδομένα της χρονοσειράς και θα δοκιμαστεί στο επόμενο 20% για προβλέψεις. Αναλυτικότερα πρόκειται για:

- Μέρες προπόνησης (Train dates) : 07/08/2015 έως 12/05/2021 (n=2106 μέρες)
- Μέρες δοκιμής πρόβλεψης (Test dates) : 13/05/2021 έως 26/09/2022(n=502 μέρες)



Διάγραμμα 5.22: Μέρες εκπαίδευσης και δοκιμής του μοντέλου Random Forest για πρόβλεψη του ETH.

Με την κλάση `ForecasterAutoreg`, δημιουργείται ένα μοντέλο και εκπαιδεύεται από έναν αναδρομητή `RandomForestRegressor` με χρονικό παράθυρο 521 καθυστερήσεων. Αυτό σημαίνει ότι το μοντέλο χρησιμοποιεί τις προηγούμενες 521 ημέρες ως βάση για την πρόβλεψη.



Διάγραμμα 5.23: Δοκιμή προβλέψεων τιμής ETH με το μοντέλο Random Forest.

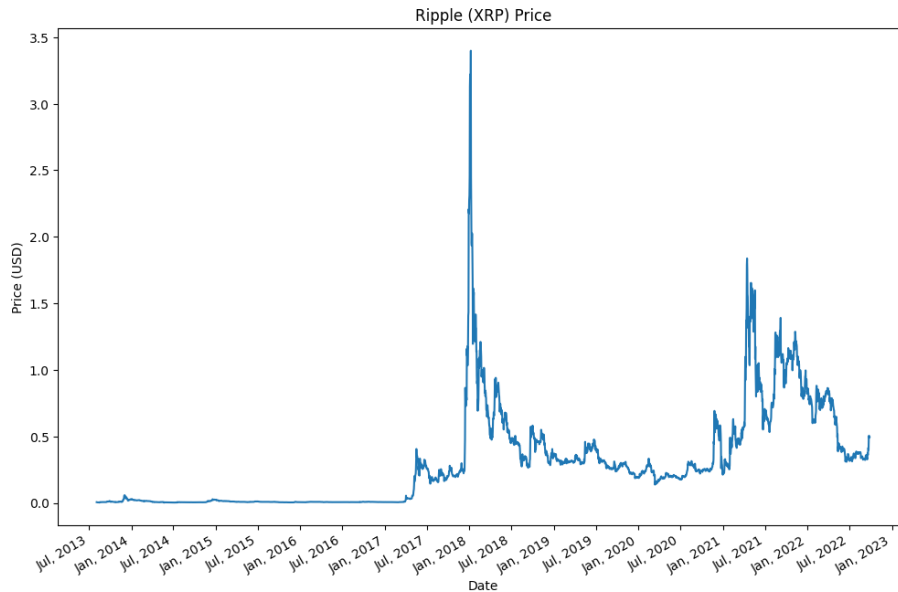
Όπως προκύπτει από το διάγραμμα 5.23 η τιμή της πρόβλεψης παραμένει σταθερή και δεν συνυπολογίζεται η τάση της χρονοσειράς. Το συγκεκριμένο μοντέλο είναι σημαντικό μόνο για βραχυχρόνιες προβλέψεις. Για να εξεταστεί το μοντέλο και να συγκριθεί με τα υπόλοιπα, απαιτείται και πάλι ο υπολογισμός του δείκτη RMSE.

Test RMSE: 1.176.

5.4 Πρόβλεψη της τιμής Ripple (XRP)

5.4.1 Μέθοδος ARIMA

Από το διάγραμμα 5.24 παρατηρείται ότι η τιμή του XRP έχει διαφορετική πορεία ανά τα χρόνια με την τιμή του BTC και του ETH. Το 2018 είναι εμφανής η ύπαρξη μέγιστης τιμής στην χρονοσειρά. Ένας νέος κύκλος εμφανίστηκε το 2021 με την μέγιστη τιμή του να συναντάται τον Ιούνιο του ίδιου έτους.



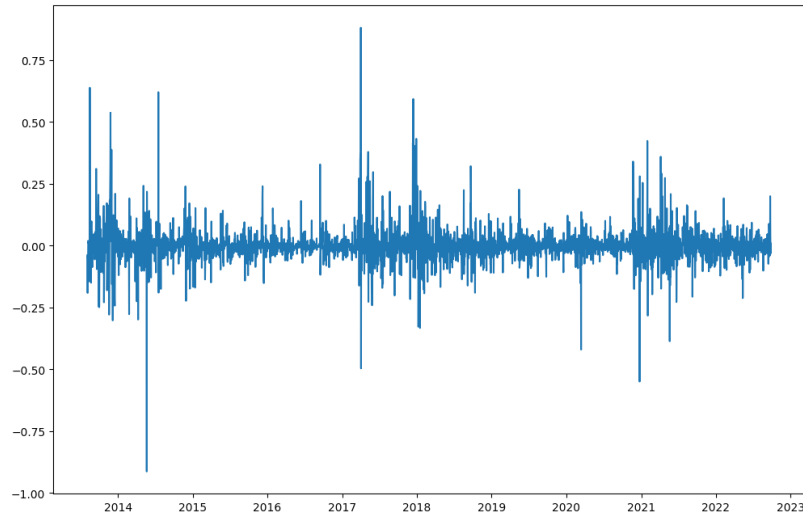
Διάγραμμα 5.24: Αναπαράσταση ιστορικών στοιχείων χρονοσειράς για την τιμή του XRP.

Στην συνέχεια είναι εμφανής η ύπαρξη τάσης, επομένως θα χρειαστούν διαφοροποιήσεις για να επιτευχθεί σταθερότητα.

Μετά από μία επανάληψη για διαφοροποίηση των δεδομένων, ο βρόγχος εξέρχεται καθώς έχει επιτύχει μια σημαντική τιμή p και εκτελούμε τη δοκιμή KPSS στα διαφορετικά δεδομένα και παρατηρούμε τα αποτελέσματα:

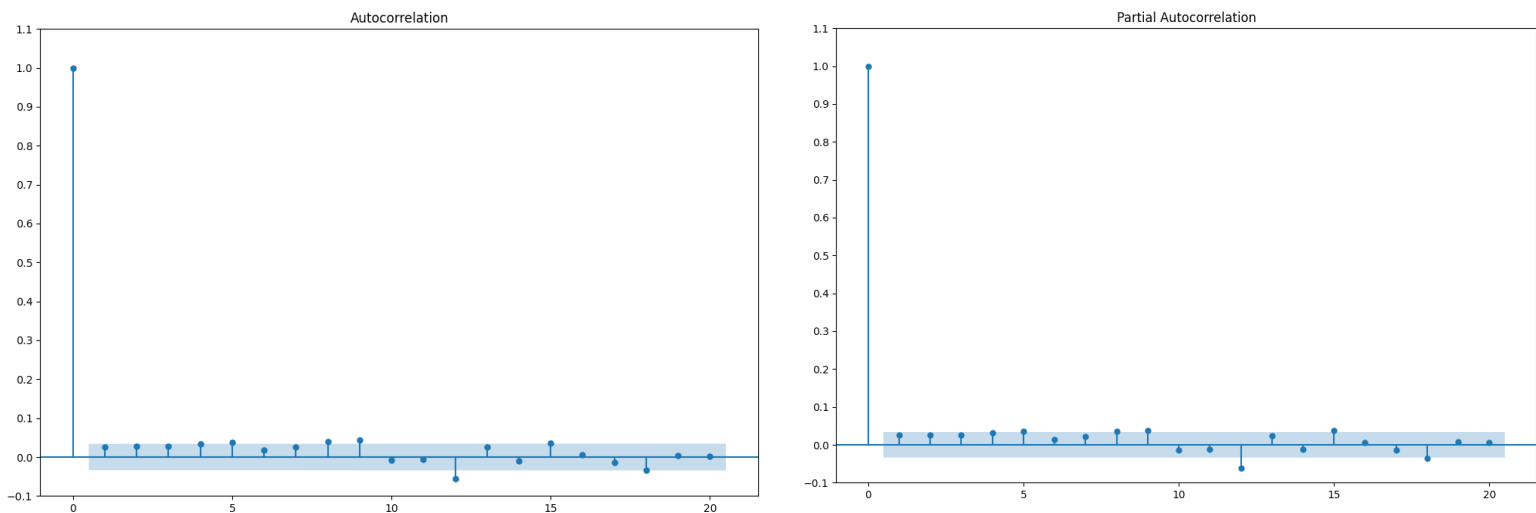
- ADF στατιστικό δοκιμής: -13.72 και p -value: 0
- KPSS στατιστικό δοκιμής: 0.057 και p -value: 0.1

Έχοντας εκτελέσει μια διαφορά πρώτης τάξης στα δεδομένα, η μηδενική υπόθεση έχει απορριφθεί για το ADF, ενώ η μηδενική υπόθεση δεν μπορεί να απορριφθεί για το KPSS, γεγονός που υποδηλώνει ότι η χρονική σειρά είναι πλέον σταθερή. Επιπλέον, αν οπτικοποιήσουμε τη σειρά, τώρα δείχνει σταθερό μέσο όρο και σταθερή διακύμανση (διάγραμμα 5.25).



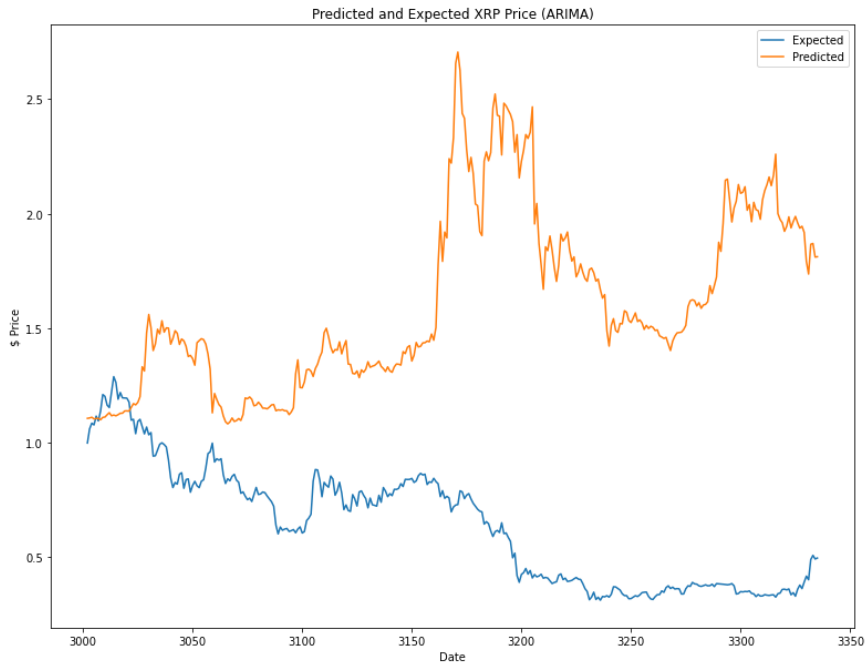
Διάγραμμα 5.25: Οπτικοποίηση δεδομένων χρονοσειράς μετά από διαφοροποίηση για επίτευξη σταθερότητας.

Στη συνέχεια, εφαρμόζουμε την μέθοδο ACF, PACF και τη δοκιμή Ljung–Box για να προσδιορίσουμε τυχόν τυχαιότητα στα δεδομένα και στα υπολείμματα του μοντέλου. Για τα δεδομένα της χρονοσειράς παίρνουμε μια τιμή p-value 0,00007, υπονοώντας ότι υπάρχουν στατιστικά στοιχεία για την απόρριψη της μηδενικής υπόθεσης άρα τα δεδομένα δεν κατανέμονται τυχαία και δεν είναι λευκός θόρυβος.



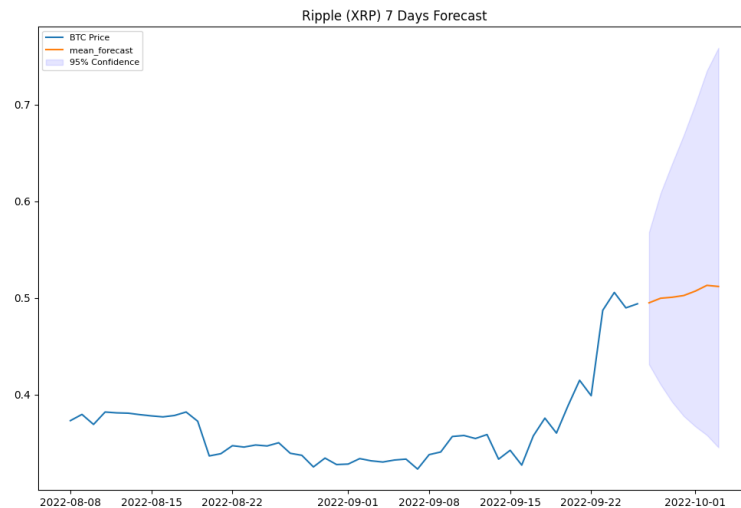
Διάγραμμα 5.26: Έλεγχος τυχαιότητας δεδομένων σε διάγραμμα ACF/ PACF.

Αφού ολοκληρώνεται ο έλεγχος σταθερότητας και τυχαιότητας, γίνεται με την βοήθεια του Auto – Arima η επιλογή του καλύτερου μοντέλου για την πρόβλεψη της συγκεκριμένης χρονοσειράς. Το καλύτερο μοντέλο είναι το ARIMA(9,1,9)(0,0,0). Αν εκπαιδεύσουμε τα δεδομένα στο 90% της χρονοσειράς και δοκιμάσουμε στο υπόλοιπο 10% τις προβλέψεις προκύπτει το διάγραμμα 5.27.



Διάγραμμα 5.27: : Τεστ του μοντέλου Πρόβλεψης της τιμής του XRP- μέθοδος ARIMA.

Διαγραμματικά, παρατηρείται ότι υπάρχει έντονη υποπρόβλεψη ενώ ταυτόχρονα το μοντέλο δεν ανταποκρίνεται ούτε στην τάση της χρονοσειράς. Το RMSE προκύπτει ίσο με 1.121.



Διάγραμμα 5.28: Πρόβλεψη της τιμής XRP για μία εβδομάδα (27/09/2022 - 4/10/2022) με διάστημα εμπιστοσύνης 95%.

Το παραπάνω γράφημα δείχνει ενδεικτικά την πρόβλεψη του μοντέλου μας της τιμής του κρυπτονομίσματος XRP για τις επόμενες επτά ημέρες και απεικονίζει το εύρος τιμών που

μπορούμε να περιμένουμε με 95% εμπιστοσύνη. Η πορτοκαλί γραμμή είναι η μέση πρόβλεψη, ενώ η σκιασμένη μωβ περιοχή έχει άνω και κάτω διαστήματα εμπιστοσύνης.

5.4.2 Μέθοδος LSTM

Η ακολουθία ορίζεται σε 100, που σημαίνει ότι το μοντέλο θα προσπαθήσει να μάθει μακροπρόθεσμες εξαρτήσεις μεταξύ χρονικών βημάτων των 100 ημερών κάθε φορά. Αυτό μπορεί να αλλάξει ανάλογα με τον πειραματισμό.

Μετά την αναμόρφωση των δεδομένων σε πίνακες 3 παραγόντων που μπορούν να τροφοδοτηθούν στο μοντέλο, τα δεδομένα χωρίζονται σε 80% εκπαίδευση και 20% δοκιμή. Αυτό αφήνει 2588 περιπτώσεις για εκπαίδευση και 648 για δοκιμή. Αυτό το μοντέλο θα έχει επίπεδα LSTM, με προεπιλογή 32 κρυφά επίπεδα. Έχει επίσης οριστεί ποσοστό εγκατάλειψης 0,2. Επιπλέον, το ποσοστό εγκατάλειψης ορίζεται σε 20%.

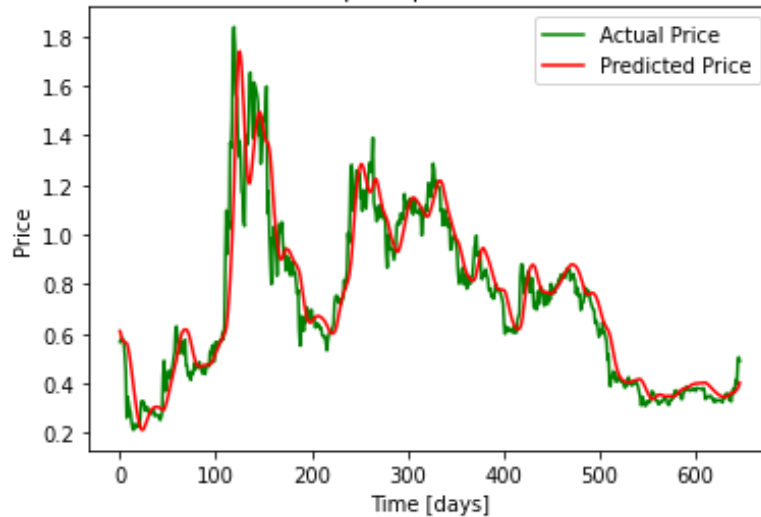
Στη συνέχεια, το μοντέλο συντάσσεται, η απώλεια μοντέλου μετράται σε MSE, το πρότυπο για γραμμικά μοντέλα για προβλέψεις. Επιλέχθηκε ο βελτιστοποιητής ADAM.

Πίνακας 5.10: Βασικό μοντέλο LSTM - πρόβλεψη XRP.

```
Model: "sequential_1"
```

Layer (type)	Output Shape	Param #
lstm_3 (LSTM)	(None, 99, 32)	4352
lstm_4 (LSTM)	(None, 99, 32)	8320
lstm_5 (LSTM)	(None, 32)	8320
dense_1 (Dense)	(None, 1)	33

Αφού επιλέχθηκε το κατάλληλο μοντέλο με την βοήθεια του βελτιστοποιητή (πίνακας 5.10), εφαρμόστηκε και δοκιμάστηκε στα δεδομένα της χρονοσειράς. Στο διάγραμμα 5.29 φαίνεται η πολύ καλή ταύτιση των προβλέψεων με τις πραγματικές τιμές. Με το παρόν μοντέλο πρόβλεψης, η τιμή του XRP θεωρείται πιθανώς προβλέψιμη.



Διάγραμμα 5.29: Τεστ του μοντέλου Πρόβλεψης της τιμής του XRP- μέθοδος LSTM.

Το RMSE προκύπτει ίσο με 0,108.

5.4.3 Μέθοδος Random Forest

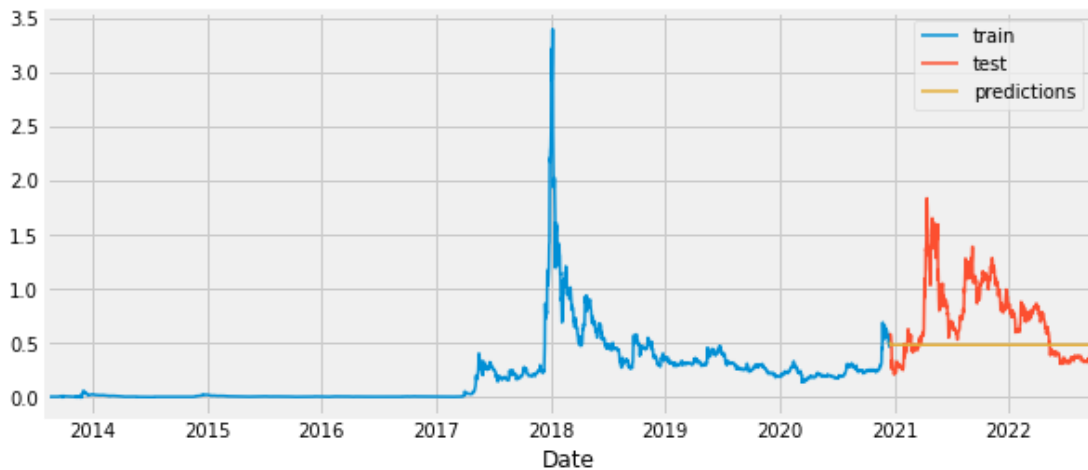
Το μοντέλο θα εκπαιδευτεί κατά 80% στα δεδομένα της χρονοσειράς και θα δοκιμαστεί στο επόμενο 20% για προβλέψεις. Αναλυτικότερα πρόκειται για:

- Μέρες προπόνησης (Train dates) : 04/08/2013 έως 17/12/2020 (n=2693 μέρες)
- Μέρες δοκιμής πρόβλεψης (Test dates) : 18/12/2020 έως 26/09/2022(n=648 μέρες)



Διάγραμμα 5.30: Μέρες εκπαίδευσης και δοκιμής του μοντέλου Random Forest για πρόβλεψη του XRP.

Με την κλάση ForecasterAutoreg, δημιουργείται ένα μοντέλο και εκπαιδεύεται από έναν αναδρομητή RandomForestRegressor με χρονικό παράθυρο 365 καθυστερήσεων. Αυτό σημαίνει ότι το μοντέλο χρησιμοποιεί τις προηγούμενες 365 ημέρες ως βάση για την πρόβλεψη.



Διάγραμμα 5.31: Δοκιμή προβλέψεων τιμής XRP με το μοντέλο Random Forest.

Όπως προκύπτει από το διάγραμμα 5.31 η τιμή της πρόβλεψης παραμένει σταθερή και δεν συνυπολογίζεται η τάση της χρονοσειράς. Το συγκεκριμένο μοντέλο είναι σημαντικό μόνο για βραχυχρόνιες προβλέψεις. Για να εξεταστεί το μοντέλο και να συγκριθεί με τα υπόλοιπα, απαιτείται και πάλι ο υπολογισμός του δείκτη RMSE.

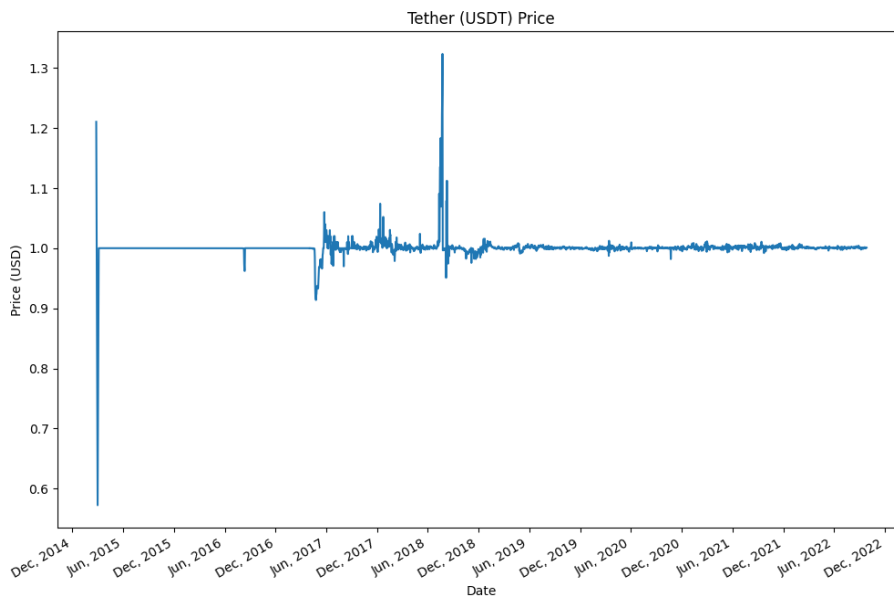
Test RMSE: 0,412.

Παρατηρείται ότι η τιμή του δείκτη RMSE είναι αρκετά μικρή. Αυτό συμβαίνει καθώς οι τιμές δεν έχουν σημαντική διακύμανση. Από αυτό το αποτέλεσμα, γίνεται κατανοητό ότι το μοντέλο πρόβλεψης για να αξιολογηθεί δεν αρκεί ένας δείκτης. Είναι απαραίτητη η διαγραμματική απεικόνιση και ο συνυπολογισμός πολλών δεικτών. Στη παρούσα χρονοσειρά, αν και το μοντέλο δεν εφαρμόζει καλά στην χρονοσειρά, η τιμή του δείκτη RMSE είναι πολύ καλή.

5.5 Πρόβλεψη της τιμής Tether (USDT)

Το κρυπτονόμισμα Tether αν και δεν καταλαμβάνει από τις υψηλότερες θέσεις στις επενδυτικές προτιμήσεις είναι αξιόλογο προς μελέτη. Εξαιτίας της σταθερής του εικόνας, κρίνεται σκόπιμο να ερευνηθεί ώστε να εξεταστούν τα προς μελέτη μοντέλα σε ένα μη ασταθές περιβάλλον και να βγουν πιο ουσιαστικά συμπεράσματα.

5.5.1 Μέθοδος ARIMA



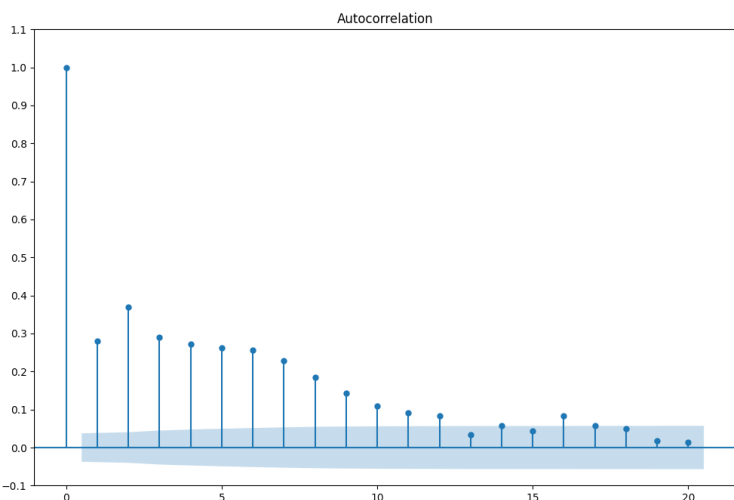
Διάγραμμα 5.32: Αναπαράσταση ιστορικών στοιχείων χρονοσειράς για την τιμή του USDT.

Από το διάγραμμα 5.32 παρατηρείται ότι η τιμή του USDT έχει σταθερή πορεία ανά τα χρόνια. Τον Ιούνιο του 2018 είναι εμφανής η ύπαρξη μεμονωμένης μέγιστης τιμής στην χρονοσειρά. Επομένως, δεν χρειάζεται κάποια διαδικασία διαφοροποίησης για να επέλθει σταθερότητα. Για την επαλήθευση των οπτικών συμπερασμάτων απαιτείται ο έλεγχος ADF και KPSS:

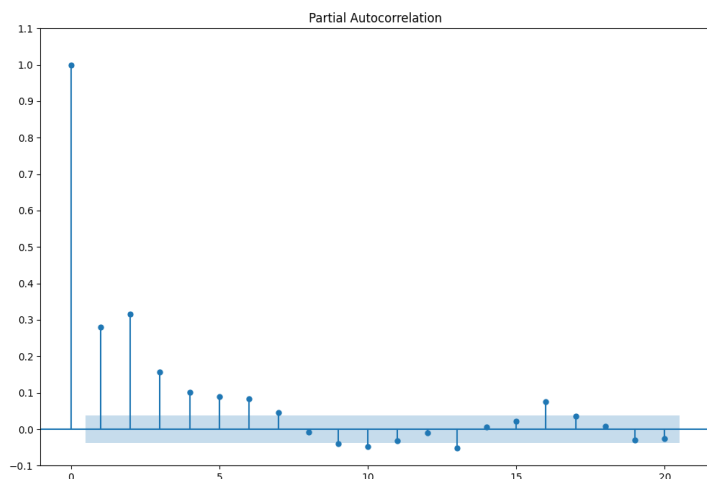
- ADF στατιστικό δοκιμής: -9.13 και p-value: 0
- KPSS στατιστικό δοκιμής: 0.089 και p-value: 0.1

Τα αποτελέσματα επαληθεύουν την παρατήρηση. Η μηδενική υπόθεση έχει απορριφθεί για το ADF, ενώ η μηδενική υπόθεση δεν μπορεί να απορριφθεί για το KPSS, γεγονός που υποδηλώνει ότι η χρονική σειρά είναι ήδη σταθερή.

Στη συνέχεια, εφαρμόζεται η δοκιμή Ljung–Box για να προσδιοριστεί πιθανή τυχαιότητα στα δεδομένα και στα υπολείμματα του μοντέλου. Για τα δεδομένα της χρονοσειράς παίρνουμε μια τιμή p-value 0, υπονοώντας ότι υπάρχουν στατιστικά στοιχεία για την απόρριψη της μηδενικής υπόθεσης άρα τα δεδομένα δεν κατανέμονται τυχαία και δεν είναι λευκός θόρυβος.

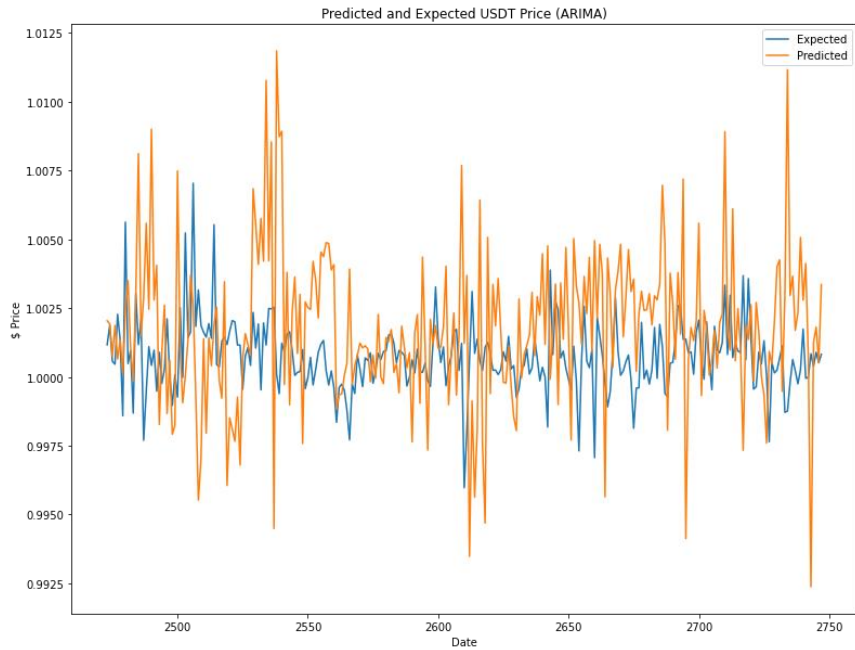


Διάγραμμα 5.33: Έλεγχος τυχειότητας δεδομένων σε διάγραμμα ACF.



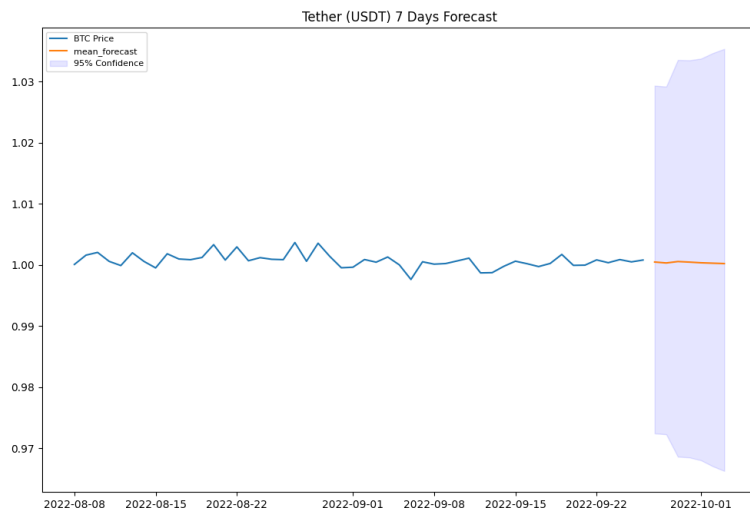
Διάγραμμα 5.34: Έλεγχος τυχειότητας δεδομένων σε διάγραμμα PACF.

Αφού ολοκληρώνεται ο έλεγχος σταθερότητας και τυχειότητας, γίνεται με την βοήθεια του Auto – Arima η επιλογή του καλύτερου μοντέλου για την πρόβλεψη της συγκεκριμένης χρονοσειράς. Το καλύτερο μοντέλο είναι το ARIMA(9,0,10)(0,0,0). Μετά την επιλογή του μοντέλου, η δοκιμή Ljung–Box στα υπολείμματα δίνει μια τιμή p 0,996, υπονοώντας ότι δεν γίνεται να απορριφθεί η μηδενική υπόθεση και τα υπολείμματά είναι λευκός θόρυβος. Επομένως, το μοντέλο εφαρμόζει καλά στα δεδομένα. Μία οπτική απεικόνιση της εφαρμογής, παρουσιάζεται στο διάγραμμα 5.35. Καθώς οι διακυμάνσεις είναι ιδιαίτερα μικρές, οι προβλέψεις δείχνουν να εφαρμόζουν καλά. Παρ'όλα αυτά είναι και πάλι εμφανή μία συντηρητική προσέγγιση του μοντέλου τόσο στα ανοδικά όσο και στα πτωτικά μέγιστα.



Διάγραμμα 5.35: Τεστ του μοντέλου Πρόβλεψης της τιμής του USDT- μέθοδος ARIMA.

Η τιμή του RMSE προκύπτει ίση με 0,003. Το παρακάτω διάγραμμα 5.36 δείχνει την πρόβλεψη του μοντέλου μας για τις επόμενες επτά ημέρες και απεικονίζει το εύρος τιμών που μπορούμε να περιμένουμε με 95% εμπιστοσύνη. Η πορτοκαλί γραμμή είναι η μέση πρόβλεψη, ενώ η σκιασμένη μωβ περιοχή έχει άνω και κάτω διαστήματα εμπιστοσύνης.



Διάγραμμα 5.36: Πρόβλεψη της τιμής USDT για μία εβδομάδα (27/09/2022 - 4/10/2022) με διάστημα εμπιστοσύνης 95%

5.5.2 Μέθοδος LSTM

Η ακολουθία ορίζεται σε 100, που σημαίνει ότι το μοντέλο θα προσπαθήσει να μάθει μακροπρόθεσμες εξαρτήσεις μεταξύ χρονικών βημάτων των 100 ημερών κάθε φορά. Αυτό μπορεί να αλλάξει ανάλογα με τον πειραματισμό.

Μετά την αναμόρφωση των δεδομένων σε πίνακες 3 παραγόντων που μπορούν να τροφοδοτηθούν στο μοντέλο, τα δεδομένα χωρίζονται σε 80% εκπαίδευση και 20% δοκιμή. Αυτό αφήνει 2118 περιπτώσεις για εκπαίδευση και 530 για δοκιμή.

Αυτό το μοντέλο θα έχει επίπεδα LSTM, με προεπιλογή 32 κρυφά επίπεδα. Έχει επίσης οριστεί ποσοστό εγκατάλειψης 0,2. Επιπλέον, το ποσοστό εγκατάλειψης ορίζεται σε 20%.

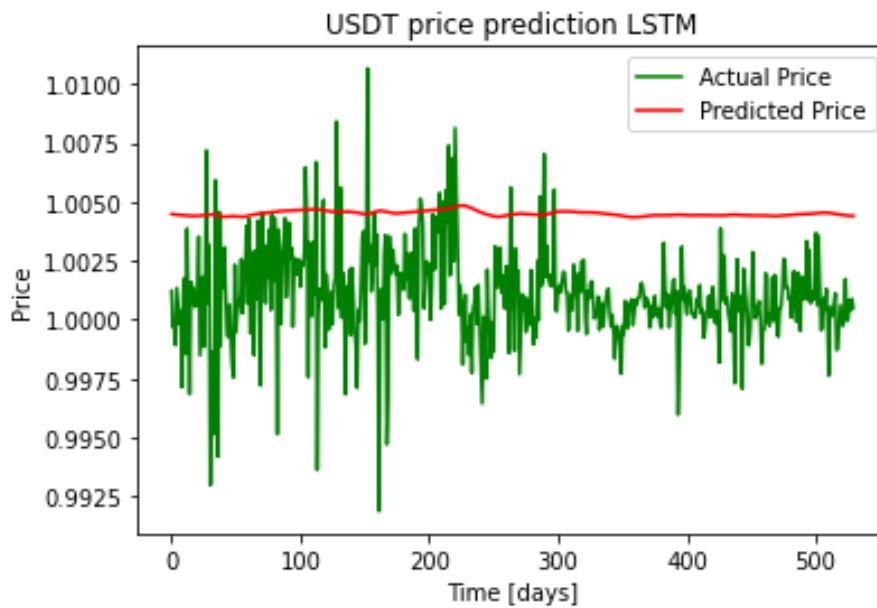
Στη συνέχεια, το μοντέλο συντάσσεται, η απώλεια μοντέλου μετράται σε MSE, το πρότυπο για γραμμικά μοντέλα για προβλέψεις. Επιλέχθηκε ο βελτιστοποιητής ADAM.

Πίνακας 5.11: : Βασικό μοντέλο LSTM - πρόβλεψη USDT.

```
Model: "sequential_3"
```

Layer (type)	Output Shape	Param #
lstm_9 (LSTM)	(None, 99, 32)	4352
lstm_10 (LSTM)	(None, 99, 32)	8320
lstm_11 (LSTM)	(None, 32)	8320
dense_3 (Dense)	(None, 1)	33

Αφού επιλεγθεί με την βοήθεια του βελτιστοποιητή το καλύτερο μοντέλο (πίνακας 5.11), οι προβλέψεις εφαρμόζονται στα πραγματικά δεδομένα της χρονοσειράς. Μία παρατήρηση στο διάγραμμα 5.37 δείχνει ότι το μοντέλο δεν είναι ικανό να εντοπίσει τις μικρές διακυμάνσεις της σειράς.



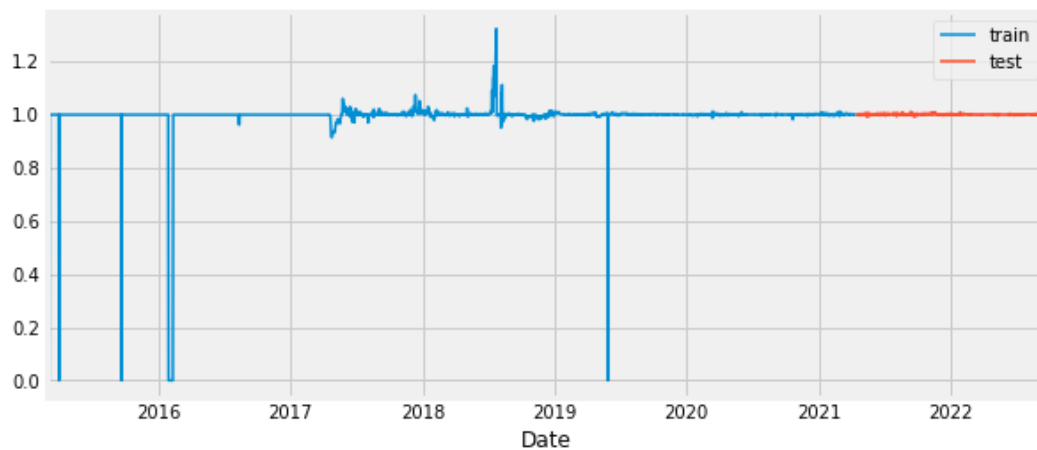
Διάγραμμα 5.37: Τεστ του μοντέλου Πρόβλεψης της τιμής του USDT- μέθοδος LSTM.

Η τιμή του RMSE με το συγκεκριμένο μοντέλο βρέθηκε ίση με 0,015.

5.5.3 Μέθοδος Random Forest

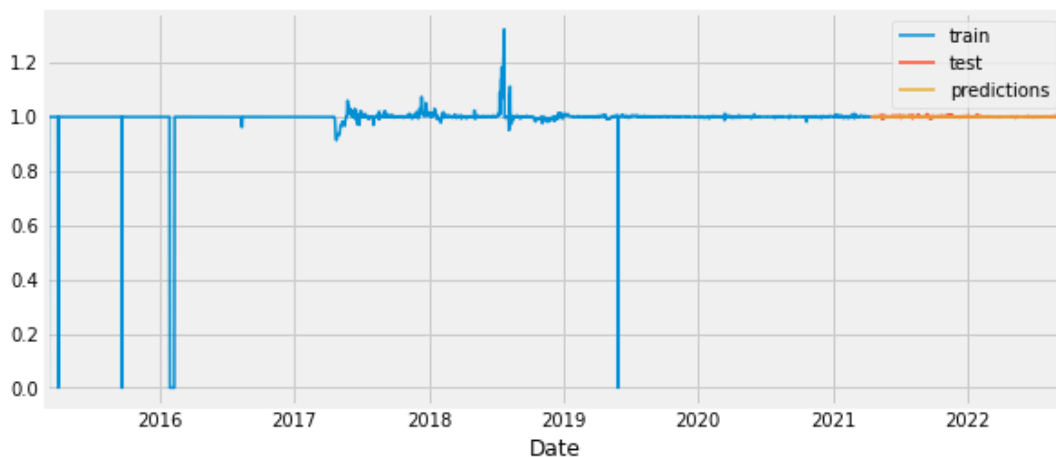
Το μοντέλο θα εκπαιδευτεί κατά 80% στα δεδομένα της χρονοσειράς και θα δοκιμαστεί στο επόμενο 20% για προβλέψεις. Αναλυτικότερα πρόκειται για:

- Μέρες προπόνησης (Train dates) : 25/02/2015 έως 14/04/2021 (n=2241 μέρες)
- Μέρες δοκιμής πρόβλεψης (Test dates) : 15/04/2020 έως 26/09/2022(n=530 μέρες)



Διάγραμμα 5.38: Μέρες εκπαίδευσης και δοκιμής του μοντέλου Random Forest για πρόβλεψη του USDT.

Με την κλάση `ForecasterAutoreg`, δημιουργείται ένα μοντέλο και εκπαιδεύεται από έναν αναδρομητή `RandomForestRegressor` με χρονικό παράθυρο 365 καθυστερήσεων. Αυτό σημαίνει ότι το μοντέλο χρησιμοποιεί τις προηγούμενες 365 ημέρες ως βάση για την πρόβλεψη.



Διάγραμμα 5.39: Δοκιμή προβλέψεων τιμής USDT με το μοντέλο Random Forest.

Στο διάγραμμα 5.39, παρατηρείται ότι οι προβλεπόμενες τιμές επαληθεύονται ακριβώς από τα δεδομένα της χρονοσειράς. Οι τιμές παραμένουν σταθερές χωρίς διακυμάνσεις. Για να εξεταστεί το μοντέλο και να συγκριθεί με τα υπόλοιπα, απαιτείται και πάλι ο υπολογισμός του δείκτη RMSE.

Test RMSE: 0,002.

5.6 Συγκέντρωση τιμών – Στατιστικός Έλεγχος

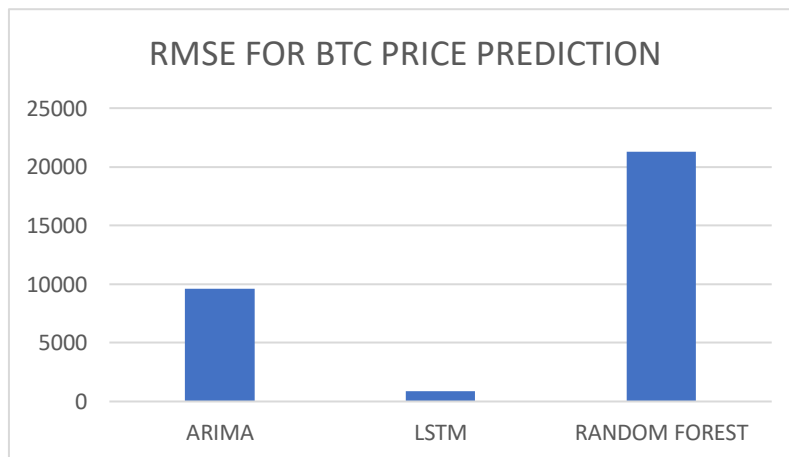
Με την βοήθεια της Python και της πλούσιας βιβλιοθήκης της έγινε μελέτη εφαρμογής μοντέλων πρόβλεψης στα κρυπτονομίσματα: Bitcoin (BTC), Ethereum (ETH), Ripple (XRP) και Tether (USDT). Τα μοντέλα πρόβλεψης που εφαρμόστηκαν για το κάθε νόμισμα ήταν το ARIMA, LSTM και το RANDOM FOREST.

Η απόδοση της πρόβλεψης αξιολογείται χρησιμοποιώντας τη μέτρηση του ριζικού μέσου τετραγώνου σφάλματος (RMSE), όπως χρησιμοποιείται ευρέως στη βιβλιογραφία επεξεργασίας σήματος και πρόβλεψης. Οι τιμές του δείκτη εμφανίζουν σημαντικές διαφορές για κάθε νόμισμα αλλά και κάθε μοντέλο. Στον πίνακα 5.12 που ακολουθεί παρουσιάζονται οι τιμές του δείκτη RMSE των προς μελέτη νομισμάτων για κάθε μοντέλο που εφαρμόστηκε.

Πίνακας 5.12: Τιμές RMSE κρυπτονομισμάτων για τα τρία μοντέλα μελέτης.

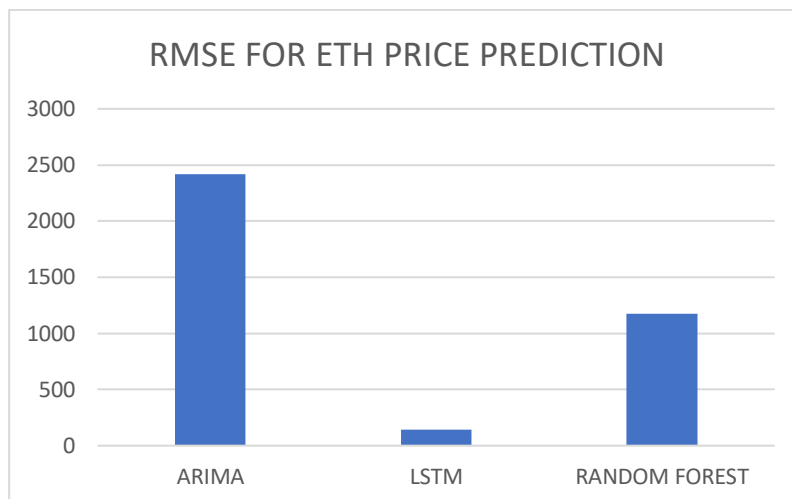
Μοντέλα	BTC	ETH	XRP	USDT
ARIMA	9.626	2.418	1.121	0,003
LSTM	866	143	0,108	0,015
RANDOM FOREST	21.280	1.176	0,412	0,002

Αναλυτικότερα, η διαγραμματική απεικόνιση του δείκτη στα ακόλουθα διαγράμματα βοηθάει στην σύγκριση της τιμής του για κάθε νόμισμα ξεχωριστά. Όσον αφορά το κορυφαίο νόμισμα BTC η τιμή του δείκτη RMSE είναι δέκα φορές μικρότερη από την αντίστοιχη με εφαρμογή του μοντέλου Arima και 20 φορές μικρότερη από την αντίστοιχη στο Random Forest (διάγραμμα 5.40).



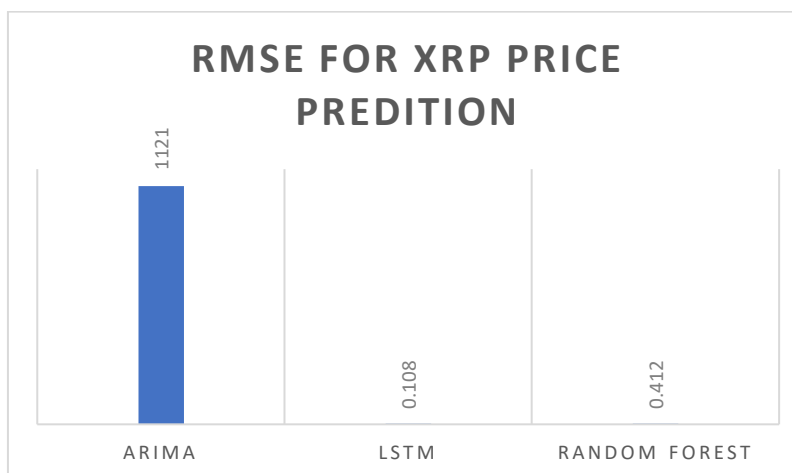
Διάγραμμα 5.40: Τιμή RMSE για πρόβλεψη του BTC με τρία διαφορετικά μοντέλα.

Στο δεύτερο δημοφιλέστερο νόμισμα Ethereum οι τιμές είναι ελαφρώς διαφοροποιημένες (διάγραμμα 5.41). Και πάλι όμως όπως και στο BTC την χαμηλότερη τιμή του δείκτη εμφανίζει το μοντέλο LSTM.



Διάγραμμα 5.41: Τιμή RMSE για πρόβλεψη του ETH με τρία διαφορετικά μοντέλα.

Η πορεία της χρονοσειράς του XRP ακολουθεί διαφορετική πορεία από τα προαναφερόμενα (διάγραμμα 5.42). Οι πιο σταθερές τιμές οδήγησαν σε καλύτερες προβλέψεις κυρίως με τα μοντέλα LSTM και Random Forest. Όμως τόσο ο δείκτης RMSE όσο και η διαγραμματική απεικόνιση των προβλέψεων στις ενότητες 5.4.2 και 5.4.3 επιβεβαιώνουν και πάλι την υπεροχή του μοντέλου LSTM.

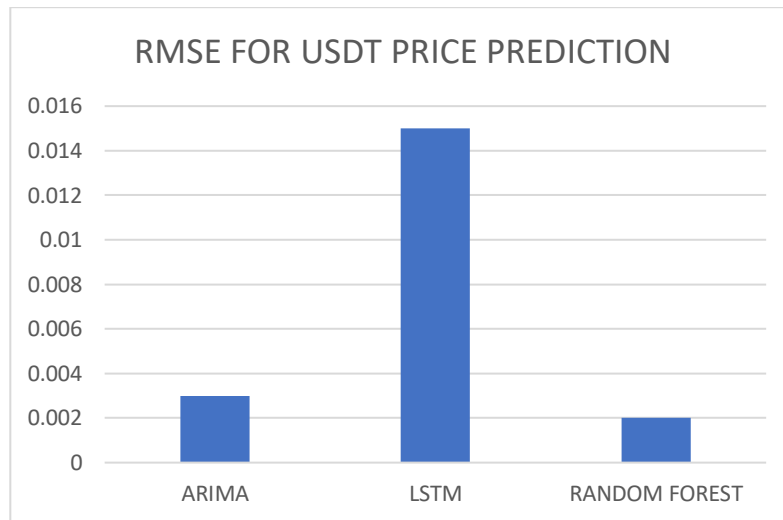


Διάγραμμα 5.42: Τιμή RMSE για πρόβλεψη του XRP με τρία διαφορετικά μοντέλα.

Τα αποτελέσματα εφαρμογής των μοντέλων στην χρονοσειρά του νομίσματος USDT είναι σαφώς διαφορετικά και οφείλονται στην σταθερότητα της τιμής του νομίσματος. Και τα τρία μοντέλα εφάρμοσαν καλά στα δεδομένα της χρονοσειράς (διάγραμμα 5.43).

Καθώς οι τιμές του δείκτη αξιολόγησης είναι ιδιαίτερος κοντά, για την καλύτερη αξιολόγηση, η παρατήρηση στο διάγραμμα εφαρμογής για κάθε μέθοδο στην χρονοσειρά είναι απαραίτητη.

Η τιμή RMSE του μοντέλου Random Forest είναι ελαφρώς καλύτερη αλλά το Arima δείχνει να εφαρμόζει καλύτερα στα δεδομένα καθώς εντοπίζει μικρές διακυμάνσεις.



Διάγραμμα 5.43: Τιμή RMSE για πρόβλεψη του USDT με τρία διαφορετικά μοντέλα.

5.7 Ανακεφαλαίωση

Η πρόβλεψη των τιμών των περισσότερων κρυπτονομισμάτων απαιτεί ιδιαίτερη προσοχή λόγω των χαρακτηριστικών δεδομένων του. Ένα ερώτημα που απασχολεί επενδυτές και ερευνητές είναι κατά πόσο αυτές οι ασταθείς και ίσως τυχαίες τιμές μπορούν να προβλεφθούν. Μπορεί η μηχανική μάθηση να βοηθήσει στην επεξεργασία τέτοιων δεδομένων;

Η μηχανική μάθηση (ML) αναφέρεται στην αυτοματοποιημένη διαδικασία μάθησης από δεδομένα χωρίς ρητές οδηγίες. Τα μοντέλα ML μπορούν να ανιχνεύσουν μη γραμμικά μοτίβα στις διακυμάνσεις των τιμών και κατά συνέπεια έχουν υιοθετηθεί ευρέως για προβλέψεις τιμών χρηματιστηρίου και κρυπτονομισμάτων. Ερευνητές παγκοσμίως έχουν χρησιμοποιήσει διάφορες τεχνικές και μεθοδολογίες από τους τομείς της στατιστικής, της μηχανικής μάθησης και της βαθιάς μάθησης, όπως LR, ARIMA, LDA, DT, RF, XGBoost, QDA, SVM και LSTM, για να προβλέψουν την τιμή του BTC και των υπόλοιπων κρυπτονομισμάτων.

Βιβλιογραφικά, υπάρχουν πολλές έρευνες και αρκετές προσπάθειες εφαρμογής των προαναφερόμενων μοντέλων πρόβλεψης για την τιμή του Bitcoin. Ερευνητικό κενό και επιδίωξη της εργασίας αποτελεί η έρευνα και άλλων ισχυρών κρυπτονομισμάτων, ώστε να

δημιουργηθούν καλύτερα συμπεράσματα για τον τρόπο πρόβλεψης της τιμής τους. Στην παρούσα εργασία έγινε προσπάθεια πρόβλεψης των τιμών τεσσάρων κρυπτονομισμάτων (Bitcoin, Ethereum, Ripple, Tether). Τα μοντέλα που χρησιμοποιήθηκαν για κάθε νόμισμα είναι το παραδοσιακό μοντέλο ARIMA, το μοντέλο χρήσης των επαναλαμβανόμενων νευρωνικών δικτύων LSTM και το μοντέλο που βασίζεται στα δένδρα αποφάσεων Random Forest. Η επιλογή των συγκεκριμένων μεθόδων έγινε για να συμπεριληφθούν τεχνικές που χρησιμοποιούν διαφορετικά πεδία, να συγκριθούν μεταξύ τους και να επιλεγεί η καλύτερη.

Με την βοήθεια της Python, επιλέχθηκε ο κατάλληλος όγκος δεδομένων για κάθε νόμισμα και ο κατάλληλος συνδυασμός παραμέτρων κάθε τεχνικής. Τα μοντέλα εκπαιδεύτηκαν σε ένα σύστημα 90/10% και 80/20% για όλα τα κρυπτονομίσματα. Για την αξιολόγηση και σύγκρισή τους χρησιμοποιήθηκε ο δείκτης RMSE.

Η εφαρμογή των μοντέλων στα δεδομένα χρονοσειράς των τιμών των Bitcoin και Ethereum ήταν παρόμοια. Οι ομοιότητες οφείλονται στην κοινή πορεία των δύο νομισμάτων στον χρόνο. Η πρόβλεψη σε αυτή την περίπτωση ήταν πολύ δύσκολη εξαιτίας των θορύβων, της τυχαιότητας και της έντονης αστάθειας. Ως εκ τούτου έγιναν πολλές αλλαγές και διαφοροποιήσεις. Το τελευταίο φάνηκε έντονα στην εφαρμογή της μεθόδου Arima, όπου για να γίνει χρήση του μοντέλου περιορίστηκε ο όγκος των δεδομένων από το 2017 και μετά. Η καλύτερη τιμή RMSE συναντάται από την εφαρμογή του μοντέλου LSTM (866 BTC και 144 ETH). Τόσο το μοντέλο Arima όσο και το μοντέλο RF δεν κατάφεραν να διεξάγουν καλές προβλέψεις. Από την άλλη, η μέθοδος LSTM φαίνεται να είναι αρκετά καλή στην πρόβλεψη της τάσης της σειράς, αλλά σταθερά κάτω από την πρόβλεψη. Το μοντέλο φαίνεται να ενεργεί συντηρητικά, υπό κυρίως της προβλέψεως. Οι καθοδικές κλίσεις φαίνεται να έχουν πιο ακριβείς προβλέψεις από τις ανοδικές τάσεις. Αν και το μοντέλο δεν φαίνεται έμπειρο στην πρόβλεψη της ίδιας της τιμής, η ικανότητά του να προβλέπει τάσεις είναι αξιοσημείωτη.

Στην περίπτωση του κρυπτονομίσματος Ripple, καλύτερη πρόβλεψη και πάλι εντοπίστηκε από το μοντέλο LSTM (RMSE: 0,108). Η διαγραμματική απεικόνιση επιβεβαιώνει την πολύ χαμηλή τιμή του δείκτη. Η δοκιμή προβλέψεων εφαρμόζει ακριβώς στα δεδομένα της χρονοσειράς. Η πρόβλεψη θεωρείται πετυχημένη στην επιλογή αυτού του μοντέλου. Αντιθέτως, το Arima και πάλι αδυνατεί να αντιμετωπίσει την αστάθεια. Επιπλέον, το RF αν και παρουσιάζει μικρό δείκτη RMSE, διαγραμματική παρατηρείται ότι δεν εντοπίζει καμία αλλαγή στην τάση.

Τέλος, η πρόβλεψη του κρυπτονομίσματος Tether αποτέλεσε μία διαφορετική διαδικασία εξαιτίας της έντονης σταθερότητάς του. Το νόμισμα αυτό εξετάστηκε κυρίως για συγκριτικούς λόγους και για να αναδειχθεί η δυσκολία προβλέψεων χωρίς σταθερότητα. Και τα τρία μοντέλα σημείωσαν ιδιαίτερα καλούς δείκτες RMSE. Διαγραμματικά, φαίνεται η υπεροχή του μοντέλου Arima, καθώς συχνά εντοπίζει τις μικρές διακυμάνσεις σε αντίθεση με τα άλλα δύο μοντέλα.

Μετά την μελέτη όλων των νομισμάτων εξαγάγεται το συμπέρασμα ότι το μοντέλο LSTM κάνει το όνειρο των ερευνητών και επενδυτών για πρόβλεψη των κρυπτονομισμάτων ρεαλιστικό. Ενώ η παραδοσιακή μέθοδος Arima και η μέθοδος δένδρων απόφασης Random Forest δεν παρείχαν ικανοποιητικές προβλέψεις, το μοντέλο LSTM εμφάνισε ελπιδοφόρα αποτελέσματα. Το LSTM μπορεί να διατηρήσει προηγούμενες πληροφορίες, οι οποίες μπορούν να βοηθήσουν σημαντικά στη βελτίωση της ικανότητάς του μοντέλου να μαθαίνει ακολουθίες σημάτων και εγγενή μη γραμμικά μοτίβα.

Κεφάλαιο 6

6 Συμπεράσματα

Κεντρικός στόχος της εργασίας αποτέλεσε η έρευνα σχετικά με την ικανότητα πρόβλεψης των κρυπτονομισμάτων. Η πρόβλεψη των τιμών των περισσότερων κρυπτονομισμάτων απαιτεί ιδιαίτερη προσοχή λόγω των χαρακτηριστικών δεδομένων του, που όπως παρατηρήθηκε είναι εξαιρετικά ασταθή, η χρονοσειρά τους εξαιρετικά μη γραμμική, μη στάσιμη, δυναμική, χωρίς καμία περιοδικότητα και με ύπαρξη φάσματος στοιχείων κλιμάκωσης. Ταυτόχρονα, εντοπίζονται θορυβώδη δεδομένα και τυχαιότητα. Ωστόσο, η παρούσα έρευνα παρουσίασε επιτυχία ως κάποιος βαθμός στη μοντελοποίηση των τιμών των κρυπτονομισμάτων με βάση διαφορετικά είδη συνόλων χαρακτηριστικών.

Αναλυτικότερα, οι τιμές του Bitcoin και του Ethereum, των δύο πιο σημαντικών κρυπτονομισμάτων, είναι στοχαστικές και κανένα δεδομένο σύνολο χαρακτηριστικών δεν μπορεί να παρέχει πλήρη πρόβλεψη. Εν τούτοις, δοκιμάστηκαν μοντέλα όπου εντοπίζουν με ακρίβεια την τάση των τιμών. Αν και τα μοντέλα δεν φαίνονται έμπειρα στην πρόβλεψη της ίδιας της τιμής των ισχυρότερων κρυπτονομισμάτων, η ικανότητά τους να προβλέπουν τάσεις είναι αξιοσημείωτη. Από την άλλη μεριά, η μελέτη λιγότερο δυναμικών κρυπτονομισμάτων ανέδειξε ότι η δημιουργία αξιόπιστων προβλέψεων είναι ένα ρεαλιστικό σενάριο.

Μετά την μελέτη των τεσσάρων κρυπτονομισμάτων (Bitcoin, Ethereum, Tether και Ripple) εξαγάγεται το συμπέρασμα ότι το μοντέλο LSTM είναι ελπιδοφόρο στον τομέα των προβλέψεων κρυπτονομισμάτων. Η εφαρμογή προηγμένων και συμβατικών αρχιτεκτονικών νευρωνικών δικτύων στα προς μελέτη διαφορετικά, ψηφιακά νομίσματα, ανέδειξε ότι το LSTM ξεπέρασε σημαντικά τα παραδοσιακά μοντέλα παλινδρόμησης, κινούμενου μέσου όρου και δένδρων αποφάσεων όσον αφορά το ριζικό μέσο τετράγωνο σφάλματος. Ο λόγος είναι ότι τα νευρωνικά συστήματα βαθιάς μάθησης είναι προφανώς αποτελεσματικά στην απομνημόνευση βραχυπρόθεσμων και μακροπρόθεσμων χρονικών πληροφοριών, γεγονός που τα καθιστά ικανά να εξάγουν κρυφά μοτίβα από τα υποκείμενα σήματα και γενικά να μοντελοποιούν χρονικές ακολουθίες, με ακρίβεια. Επίσης, τα LSTMs βρέθηκαν σημαντικά πιο επιτυχημένα στην εκμάθηση μοτίβων για το δείγμα δοκιμής, το οποίο ουσιαστικά περιελάμβανε την περίοδο πρόβλεψης.

Βιβλιογραφία

Ελληνική

Κριάρης, Π. (2022, June 26). Κρυπτονομίσματα: τι είναι και γιατί πέφτουν. *ΚΑΘΗΜΕΡΙΝΗ*.
<https://www.kathimerini.gr/economy/561927424/kryptonomismata-ti-einai-kai-giati-peftoun/>

Ξένα

Akyildirim, E., Goncu, A., & Sensoy, A. (2020). Prediction of cryptocurrency returns using machine learning. *Annals of Operations Research*, 297(1–2), 3–36.
<https://doi.org/10.1007/s10479-020-03575-y>

Alaoui, M. El, Bouri, E., & Roubaud, D. (2019). Bitcoin price–volume: A multifractal cross-correlation approach. *Finance Research Letters*, 31(November 2018), 374–381.
<https://doi.org/10.1016/j.frl.2018.12.011>

Breiman, L. (2001a). Random Forests. In R. E. Schapire (Ed.), *Kluwer Academic Publishers: Vol. 12343 LNCS*. Kluwer Academic Publishers. https://doi.org/10.1007/978-3-030-62008-0_35

Breiman, L. (2001b). Statistical modeling: The two cultures. *Statistical Science*, 16(3), 199–215. <https://doi.org/10.1214/ss/1009213726>

Burke, J. J. (2021). Financial Services in the Twenty-First Century. In *Financial Services in the Twenty-First Century*. <https://doi.org/10.1007/978-3-030-63967-9>

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3), 1247–1250. <https://doi.org/10.5194/gmd-7-1247-2014>

Chowdhury, R., Rahman, M. A., Rahman, M. S., & Mahdy, M. R. C. (2020). An approach to predict and forecast the price of constituents and index of cryptocurrency using machine learning. *Physica A: Statistical Mechanics and Its Applications*, 551, 124569.

<https://doi.org/10.1016/j.physa.2020.124569>

Delahaye, J.-P. (2016). Cryptocurrencies and Blockchains. *Inference: International Review of Science*, 2(4). <https://doi.org/10.37282/991819.16.38>

Greaves, A., & Au, B. (2015). Using the bitcoin transaction graph to predict the price of bitcoin. *Quoted*, 3, 22. http://snap.stanford.edu/class/cs224w-2015/projects_2015/Using_the_Bitcoin_Transaction_Graph_to_Predict_the_Price_of_Bitcoin.pdf

Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., & Schmidhuber, J. (2017). LSTM: A Search Space Odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, 28(10), 2222–2232. <https://doi.org/10.1109/TNNLS.2016.2582924>

Hamayel, M. J., & Owda, A. Y. (2021). A Novel Cryptocurrency Price Prediction Model Using GRU, LSTM and bi-LSTM Machine Learning Algorithms. *Ai*, 2(4), 477–496. <https://doi.org/10.3390/ai2040030>

Jang, H., & Lee, J. (2018). An Empirical Study on Modeling and Prediction of Bitcoin Prices With Bayesian Neural Networks Based on Blockchain Information. *IEEE Access*, 6, 5427–5437. <https://doi.org/10.1109/ACCESS.2017.2779181>

Kane, M. J., Price, N., Scotch, M., & Rabinowitz, P. (2014). Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks. *BMC Bioinformatics*, 15(1). <https://doi.org/10.1186/1471-2105-15-276>

Khedhiri, S., & Montasser, G. El. (2010). The effects of additive outliers on the seasonal KPSS test: A Monte Carlo analysis. *Journal of Statistical Computation and Simulation*, 80(6), 643–651. <https://doi.org/10.1080/00949650902755160>

Kim, S., & Kim, H. (2016). A new metric of absolute percentage error for intermittent demand forecasts. *International Journal of Forecasting*, 32(3), 669–679. <https://doi.org/10.1016/j.ijforecast.2015.12.003>

Lahmiri, S., & Bekiros, S. (2019). Cryptocurrency forecasting with deep learning chaotic neural networks. *Chaos, Solitons and Fractals*, 118, 35–40.

<https://doi.org/10.1016/j.chaos.2018.11.014>

Lopez, J. H. (1997). The power of the ADF test. *Economics Letters*, 57(1), 5–10.

[https://doi.org/10.1016/s0165-1765\(97\)81872-1](https://doi.org/10.1016/s0165-1765(97)81872-1)

Makarov, I., & Schoar, A. (2021). BLOCKCHAIN ANALYSIS OF THE BITCOIN MARKET. *National Bureau of Economic Research (NBER)*, 7, 6. <https://doi.org/10.3386/w29396>

McNally, S., Roche, J., & Caton, S. (2018). Predicting the Price of Bitcoin Using Machine Learning. *26th Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP)*, 339–343.

<https://doi.org/10.1109/PDP2018.2018.00060>

Natarajan, H., Krause, S. K., & Gradstein, H. L. (2017). Distributed Ledger Technology (DLT) and Blockchain. *FinTech Note*, 1, 1–60.

<http://hdl.handle.net/10986/29053%0Ahttp://documents.worldbank.org/curated/en/177911513714062215/pdf/122140-WP-PUBLIC-Distributed-Ledger-Technology-and-Blockchain-Fintech-Notes.pdf>

Prayudani, S., Hizriadi, A., Lase, Y. Y., Fatmi, Y., & Al-Khowarizmi. (2019). Analysis Accuracy of Forecasting Measurement Technique on Random K-Nearest Neighbor (RKNN) Using MAPE and MSE. *Journal of Physics: Conference Series*, 1361(1).

<https://doi.org/10.1088/1742-6596/1361/1/012089>

Rhif, M., Abbes, A. Ben, Farah, I. R., Martínez, B., & Sang, Y. (2019). Wavelet transform application for/in non-stationary time-series analysis: A review. *Applied Sciences (Switzerland)*, 9(7), 1–22. <https://doi.org/10.3390/app9071345>

Roggero, M. (2012). Discontinuity detection and removal from data time series. *VII Hotine-Marussi Symposium*, 137(January 2012). <https://doi.org/10.1007/978-3-642-22078-4>

Satoshi, N. (2017). Bitcoin: A Peer-to-Peer Electronic Cash System. *Artificial Life*, 23(4), 552–557. https://doi.org/10.1162/ARTL_a_00247

Schurman, K. (2012). *Bitcoin: Free Money or Fraud?* Hyperink, Marston Gate.

<https://books.google.gr/books?id=GEx->

BwAAQBAJ&lpg=PT47&ots=XY1mx0tFwG&dq=gawker and
btc&hl=el&pg=PT47#v=onepage&q&f=false

Sherratt, F., Plummer, A., & Iravani, P. (2021). Understanding lstm network behaviour of imu-based locomotion mode recognition for applications in prostheses and wearables. *Sensors (Switzerland)*, 21(4), 1–23. <https://doi.org/10.3390/s21041264>

Witzig, P., & Salomon, V. (2019). Cutting out the middleman: A case study of blockchain technology induced reconfigurations in the Swiss financial services industry. *Blockchains, Smart Contracts, Decentralised Autonomous Organisations and the Law*, 18–48. <https://doi.org/10.4337/9781788115131.00008>

Διαδικτυακοί τόποι

CoinLoan. (2020). *The complete history of Litecoin*. www.coinloan.io.
<https://coinloan.io/article/the-complete-history-of-litecoin-ltc/>

Frankenfield, J. (2022). *Cryptocurrency Explained With Pros and Cons for Investment*. www.investopedia.com. <https://www.investopedia.com/terms/c/cryptocurrency.asp>

Hamacher, A. (2021). *What Are Flash Loans? The DeFi Lending Phenomenon Explained*. www.decrypt.co. <https://decrypt.co/resources/what-are-flash-loans-the-defi-lending-phenomenon-explained>

Hayes, A. (2022). *Who Is Satoshi Nakamoto?* www.investopedia.com.
<https://www.investopedia.com/terms/s/satoshi-nakamoto.asp>

Hyndman, R. J. (2011). *Cyclic and seasonal time series*.
<https://robjhyndman.com/hyndsight/cyclicts/>

Jones, E. (2021). *A brief history of Cryptocurrency*. www.cryptovantage.com.
<https://www.cryptovantage.com/guides/a-brief-history-of-cryptocurrency/>

Kriptomat. (2022). *Τα Πιο Δημοφιλή Δίκτυα Blockchain*.
<https://kriptomat.io/gr/blockchain/ta-pio-dimofili-diktya-blockchain/>

Marr, B. (2021). *Blockchain: A Very Short History Of Ethereum Everyone Should Read*.

www.bernardmarr.com. <https://bernardmarr.com/blockchain-a-very-short-history-of-ethereum-everyone-should-read/>

Myre, G. (2021). *How Bitcoin Has Fueled Ransomware Attacks*. www.npr.org.
<https://www.npr.org/2021/06/10/1004874311/how-bitcoin-has-fueled-ransomware-attacks>

Palachy, S. (2019). Detecting stationarity in time series data. *Towards Data Science*.
<https://towardsdatascience.com/detecting-stationarity-in-time-series-data-d29e0a21e638>

Peixeiro, M. (2019). *The Complete Guide to Time Series Analysis and Forecasting*. Towards Data Science. <https://towardsdatascience.com/the-complete-guide-to-time-series-analysis-and-forecasting-70d476bfe775>

Perloth, N., Griffith, E., & Benner, K. (2021). Pipeline Investigation Upends Idea That Bitcoin Is Untraceable. *The New York Times*.
<https://www.nytimes.com/2021/06/09/technology/bitcoin-untraceable-pipeline-ransomware.html>

Redman, J. (2019). *Hackers Have Looted More Bitcoin Than Satoshi's Entire Stash*. News.Bitcoin.Com. <https://news.bitcoin.com/hackers-have-looted-more-bitcoin-than-satoshis-entire-stash/>

Roggero, M. (2012). Discontinuity detection and removal from data time series. *VII Hotine-Marussi Symposium*, 137(January 2012). <https://doi.org/10.1007/978-3-642-22078-4>

Satoshi, N. (2017). Bitcoin: A Peer-to-Peer Electronic Cash System. *Artificial Life*, 23(4), 552–557. https://doi.org/10.1162/ARTL_a_00247

toppr.com. (2020). *Time Series Analysis*. <https://www.toppr.com/guides/business-mathematics-and-statistics/time-series-analysis/components-of-time-series/>

Tretina, K., & Powell, F. (2022). *Top 10 Cryptocurrencies Of 2022*. www.forbes.com.
<https://www.forbes.com/advisor/investing/cryptocurrency/top-10-cryptocurrencies/>

Tuwiner, J. (2022). *Mt Gox Hack Explained*. Wallabit Media LLC.

<https://buybitcoinworldwide.com/mt-gox-hack/>

www.coinmarketcap.com. (n.d.). *Currencies - BitCoin*. CoinMarketCap. Retrieved May 7, 2022, from <https://coinmarketcap.com/currencies/bitcoin/>