



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ ΔΙΑΤΑΞΕΩΝ ΚΑΙ
ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ

**Συσχέτιση και πρόβλεψη τιμών κρυπτονομισμάτων με
ομαδοποίηση βάσει των τεχνολογιών τους**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ιωάννης, Εμμανουήλ, Γκιορτζής

Επιβλέπων : Βασίλειος Ασημακόπουλος
Ομότιμος Καθηγητής Ε.Μ.Π

Υπεύθυνος : Ευάγγελος Σπηλιώτης
Διδάκτωρ Ε.Μ.Π

Αθήνα, Οκτώβριος, 2023



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ ΔΙΑΤΑΞΕΩΝ ΚΑΙ
ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ

Συσχέτιση και πρόβλεψη τιμών κρυπτονομισμάτων με ομαδοποίηση βάσει των τεχνολογιών τους

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ιωάννης, Εμμανουήλ, Γκιιορτζής

Επιβλέπων : Βασίλειος Ασημακόπουλος

Ομότιμος Καθηγητής Ε.Μ.Π

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την.

.....
Β. Ασημακόπουλος,
Ομότιμος Καθηγητής Ε.Μ.Π

.....
Χ. Δούκας,
Καθηγητής Ε.Μ.Π

.....
Δ. Ασκούνης,
Καθηγητής Ε.Μ.Π

Αθήνα, Οκτώβριος, 2023



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ ΔΙΑΤΑΞΕΩΝ ΚΑΙ
ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ

.....

ΙΩΑΝΝΗΣ ΓΚΙΟΡΤΖΗΣ

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών ΕΜΠ.

Copyright © 2023 ΙΩΑΝΝΗΣ ΓΚΙΟΡΤΖΗΣ, 2023

Με επιφύλαξη παντός δικαιώματος. All rights reserved

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Η παρούσα διπλωματική εργασία αποσκοπεί στην ανάπτυξη μιας γενικότερης μεθοδολογίας πρόβλεψης της τιμής διαφόρων κρυπτονομισμάτων βάσει των τεχνολογικών χαρακτηριστικών τους, καθώς και την αξιολόγηση και ερμηνεία της σχέσης μεταξύ των τεχνολογιών αυτών. Το εν λόγω τεχνολογικό χαρακτηριστικό που μελετάται είναι ο αλγόριθμος ομοφωνίας των κρυπτονομισμάτων, ο οποίος συνθέτει τους όρους για την επαλήθευση μίας συναλλαγής. Αναλόγως του αλγορίθμου που χρησιμοποιείται, γεννώνται ειδοποιείς διαφορές που επηρεάζουν την δημοτικότητα και χρήση των κρυπτονομισμάτων. Οι αλγόριθμοι αυτοί θα καθορίσουν τις κατηγορίες στις οποίες θα χωριστούν τα δεδομένα εισόδου.

Η βασική ιδέα της προτεινόμενης μεθοδολογίας είναι η εξάλειψη της αστάθειας που παρουσιάζουν οι τιμές των κρυπτονομισμάτων, χρησιμοποιώντας τα σύνολα των ομάδων ως παράγοντες εξομάλυνσης της τιμής κατά την διαδικασία της πρόβλεψης. Το προτεινόμενο μοντέλο δέχεται ως είσοδο το σύνολο ιστορικών τιμών επιλεγμένων κρυπτονομισμάτων, του ίδιου αλγορίθμου ομοφωνίας, και παράγει ως έξοδο την πρόβλεψη για ένα συγκεκριμένο κρυπτονόμισμα στόχο, διαφορετικού αλγορίθμου ομοφωνίας. Με βάση τις προβλέψεις, γίνεται προσπάθεια, να αναλυθεί η σχέση μεταξύ των τεχνολογιών και το πως η μία επηρεάζει την άλλη. Τα δεδομένα εισόδου αποτελούν τέσσερις χρονοσειρές από ομάδες διαφορετικών αλγορίθμων ομοφωνίας, όπου κάθε μία συντίθεται από τον σταθμισμένο μέσο όρο της τιμής, των επιλεγμένων κρυπτονομισμάτων, επί του χρηματιστηριακού τους όγκου. Κάθε χρονοσειρά τροφοδοτείται στο μοντέλο και χρησιμοποιείται για την πρόβλεψη τόσο των άλλων αλγορίθμων ομοφωνίας, όσο και των μεμονωμένων κρυπτονομισμάτων.

Για την δημιουργία του μοντέλου χρησιμοποιήθηκαν ιστορικά δεδομένα που αντλήθηκαν από το Yahoo Finance για κάθε κρυπτονόμισμα και το μοντέλο υλοποιήθηκε στην γλώσσα προγραμματισμού Python. Οι προβλέψεις αξιολογήθηκαν με γνώμονα τους δείκτες ακρίβειας και τελικά επιλέχθηκε το μοντέλο με τα λιγότερα σφάλματα. Το βέλτιστο μοντέλο αξιολογήθηκε περαιτέρω προβλέποντας την πτώση ή άνοδο του κάθε κρυπτονομίσματος, προσομοιώνοντας την χρήση του σαν επενδυτικό εργαλείο. Τα τελικά αποτελέσματα συγκρίθηκαν με απλές χρηματιστηριακές στρατηγικές σε σύγκριση με το κέρδος που επέφεραν. Σε σύγκριση με την πιο απλή μέθοδο επένδυσης, το μοντέλο έβγαλε πολύ καλύτερα αποτελέσματα, έχοντας πάντα επενδυτικό κέρδος, ενώ σε σύγκριση με την βέλτιστη στρατηγική, το μοντέλο παρουσίασε ικανοποιητικά αποτελέσματα, πλησιάζοντας σε κάποιες περιπτώσεις και το κέρδος της βέλτιστης επένδυσης.

Στο πλαίσιο της παρούσας διπλωματικής εργασίας περιγράφεται αναλυτικά η πειραματική αυτή διαδικασία και παρουσιάζονται τα αποτελέσματα και τα συμπεράσματα που προέκυψαν.

Λέξεις Κλειδιά: Bitcoin, Κρυπτονομίσματα, Τεχνικές Προβλέψεων, Χρονοσειρές

Abstract

This diploma thesis aims to develop a general methodology for accurate prediction of different cryptocurrency prices, based on their underlying technologies, as well as to analyze the commonalities of the aforementioned technologies. The technology in question is the consensus algorithm of cryptocurrencies, which consists of the building blocks for validating the blockchain transaction. Depending on the algorithm used, noteworthy differences come up, that affect the popularity and use of cryptocurrencies. These algorithms will determine the categories into which the input data will be divided.

The basic hypothesis of the proposed methodology is the elimination of cryptocurrency instability, in price prediction, using a combination of cryptocurrencies with the same underlying technology, as a smoothing factor. The proposed model is provided with the sum of historical cryptocurrency data of the same consensus algorithm as input, and generates a prediction for another cryptocurrency, using different consensus. Based on the resulting predictions, an analysis of the connection between the different technologies is made. The input data consists of four time-series, each one representing a different consensus algorithm, created by the weighted sum of the cryptocurrency price times its volume. Each time-series is fed to the model to predict the price of not only the other time series, but also the price of the isolated cryptocurrencies.

For the creation of the model, all the historical data for every cryptocurrency was extracted from Yahoo Finance and the prediction algorithm was written in the Python programming language. The predictions were evaluated in terms of accuracy indices and finally the model with the least errors was selected. The optimal model was further evaluated by predicting the fall or rise of each cryptocurrency, simulating its use as an investment tool. The results were compared to simple stock market strategies compared to the profit they produced. Compared to the simplest investment method, the model produced much better results, always having an investment profit, while compared to the optimal strategy, the model showed satisfactory results, in some cases even approaching the optimal profit.

Within this thesis, this experimental procedure is described in detail and the results and conclusions obtained are presented.

Keywords: Bitcoin, Cryptocurrencies, Forecasting Techniques, Time-series

Ευχαριστίες

Η παρούσα διπλωματική εργασία εκπονήθηκε στα πλαίσια των ερευνητικών δραστηριοτήτων της Μονάδας Προβλέψεων και Στρατηγικής κατά το ακαδημαϊκό έτος 2022 – 2023. Η μονάδα υπάγεται στον Τομέα Βιομηχανικών Διατάξεων και Συστημάτων Αποφάσεων της Σχολής Ηλεκτρολόγων Μηχανικών & Μηχανικών Η/Υ, του Εθνικού Μετσόβιου Πολυτεχνείου.

Αρχικά θα ήθελα να ευχαριστήσω τον Ομότιμο Καθηγητή κ. Βασίλειο Ασημακόπουλο για την ανάθεση της παρούσας εργασίας και την ευκαιρία που μου έδωσε να εντρυφήσω στον κλάδο των τεχνικών προβλέψεων και συγκεκριμένα στις προβλέψεις κρυπτονομισμάτων. Θα ήθελα ακόμα να ευχαριστήσω τον Καθηγητή κ. Χρυσόστομο Δούκα και τον Καθηγητή κ. Δημήτριο Ασκούνη για τη συμμετοχή τους στην επιτροπή εξέτασης της εργασίας. Ακόμα θα ήθελα να ευχαριστήσω θερμά τον Δρ. Ευάγγελο Σπηλιώτη, για τη συνεισφορά και την καθοδήγηση σε όλη την πρόοδο της παρούσας διπλωματικής εργασίας.

Τέλος θα ήθελα να ευχαριστήσω την οικογένεια μου και τους φίλους μου για την υποστήριξη τους καθ' όλη τη διάρκεια των σπουδών μου.

Γκιορτζής Ιωάννης,
Αθήνα, Οκτώβριος 2023

Περιεχόμενα

Κατάλογος σχημάτων	iv
Κατάλογος Πινάκων	vii
Κεφάλαιο 1. Εισαγωγή	1
1.1 Αντικείμενο εργασίας	1
1.2 Οργάνωση εργασίας.....	2
Κεφάλαιο 2. Μελέτη των κρυπτονομισμάτων	4
2.1 Βασικές Έννοιες	4
2.2 Η έννοια του Blockchain.....	4
2.3 Εφαρμογές του Blockchain.....	6
2.4 Τεχνολογία των κρυπτονομισμάτων	7
2.5 Βασικές κατηγορίες κρυπτονομισμάτων βάσει αλγορίθμου ομοφωνίας.....	8
2.5.1 Απόδειξη εργασίας (Proof-of-Work)	8
2.5.2 Απόδειξη κυριότητας (Proof-of-Stake)	9
2.5.3 Απόδειξη Εξουσίας (Proof-of-Authority)	9
2.5.4 Απόδειξη Χωρητικότητας (Proof-of-Capacity)	9
2.5.5 Υβριδικοί Αλγόριθμοι (Hybrid Algorithms)	10
2.5.6 Αλγόριθμος αντοχής σε βυζαντινά σφάλματα (Byzantine Fault Tolerance - BFT)	10
Κεφάλαιο 3. Ιστορική αναδρομή και βιβλιογραφική ανασκόπηση	11
3.1 Ιστορική αναδρομή.....	11
3.2 Τα κυριότερα κρυπτονομίσματα	12
3.2.1 Bitcoin	12
3.2.2 Litecoin	15
3.2.3 Ethereum	16
3.3 Εμπορία & ανταλλαγή κρυπτονομισμάτων (Cryptocurrency trading)	18
3.4 Επένδυση σε κρυπτονομίσματα	20
3.5 Σύντομη βιβλιογραφική επισκόπηση	21
Κεφάλαιο 4. Μέθοδοι προβλέψεων	24
4.1 Είδη προβλέψεων.....	24
4.1.1 Ποιοτικές μέθοδοι	24
4.1.2 Ποσοτικές Μέθοδοι.....	25
4.2 Χρονοσειρές και ποιοτικά τους χαρακτηριστικά.....	25
4.3 Μοντέλο χρονοσειρών.....	27
4.3.1 Αφελής μέθοδος (Naive).....	27

4.3.2	Εποχική αφελής μέθοδος (Seasonal Naive).....	27
4.3.3	Μέθοδος κινητού μέσου όρου (ΚΜΟ).....	28
4.3.4	Μέθοδος εκθετικής εξομάλυνσης	28
4.3.5	Αυτοπαλινδρομικά μοντέλα κινητού μέσου όρου (ARIMA).....	31
4.4	Διαιτιοκρατικό μοντέλο	33
4.4.1	Απλή γραμμική παλινδρόμηση.....	34
4.4.2	Δειγματικός συντελεστής συσχέτισης (Pearson).....	36
4.4.3	Σχέση του συντελεστή συσχέτισης και παλινδρόμησης	36
4.4.4	Στατιστικοί δείκτες.....	37
4.4.5	Πολλαπλή γραμμική παλινδρόμηση	40
4.5	Δείκτες Ακρίβειας	40
4.5.1	Σφάλματα εξαρτώμενα από την κλίμακα.....	40
4.5.2	Ποσοστιαία σφάλματα	42
4.5.3	Σχετικά σφάλματα	43
Κεφάλαιο 5. Πειραματική διαδικασία		45
5.1	Συλλογή δεδομένων	45
5.2	Σύγκριση των αρχικών δεδομένων.....	48
5.3	Επεξεργασία δεδομένων και δημιουργία γενικών δεικτών αλγορίθμων ομοφωνίας ..	50
Κεφάλαιο 6. Μελέτη του μοντέλου και ανάλυση αποτελεσμάτων		55
6.1	Διαχωρισμός δεδομένων	55
6.2	Επιλογή του βέλτιστου μοντέλου	57
6.2.1	Μοντέλο με χρήση βέλτιστων ολισθήσεων.....	57
6.2.2	Μοντέλο με χρήση βέλτιστων ολισθήσεων και ολισθήσεων κατά μία ημέρα	59
6.2.3	Μοντέλο με χρήση βέλτιστων ολισθήσεων, ολισθήσεων κατά μία ημέρα και ολισθήσεων γειτονικών όρων επιλεγμένων χειροκίνητα (cherry-picked)	61
6.2.4	Μοντέλο καλύτερων ολισθήσεων με επανεκπαίδευση.....	62
6.2.5	Μοντέλο καλύτερων ολισθήσεων με επανεκπαίδευση και αφαίρεση των παλαιότερων δεδομένων.....	64
6.2.6	Βέλτιστο μοντέλο με προσθήκη του δείκτη S&P500.....	65
6.2.7	Βέλτιστο μοντέλο με προσθήκη του δείκτη του αργού πετρελαίου	67
6.2.8	Βέλτιστο μοντέλο με προσθήκη του δείκτη του χρυσού.....	68
6.3	Αξιολόγηση αποτελεσμάτων και πρόβλεψη τιμής	69
6.4	Πρόβλεψη τιμών μεμονωμένων κρυπτονομισμάτων.....	71
Κεφάλαιο 7. Επίλογος.....		73
7.1	Σύνοψη και συμπεράσματα	73
7.2	Μελλοντικές προεκτάσεις.....	75
7.2.1	Έλεγχος ακρίβειας με χρήση μη γραμμικών μοντέλων	75

7.2.2	Έλεγχος ακρίβειας σε συναλλαγές πραγματικού χρόνου (real time trading)	75
7.2.3	Δημιουργία πληροφοριακού συστήματος που εκτελεί συναλλαγές με χρήση του μοντέλου	75
7.2.4	Διεύρυνση και μελέτη βέλτιστων ολισθήσεων (ορίζοντας πρόβλεψης).....	76
Βιβλιογραφία.....		77

Κατάλογος σχημάτων

Εικόνα 1: Απεικόνιση ιεραρχίας αλυσίδας block. Dzulfikar, F. & Susanto, A., Simplified Bitcoin Block Chain (2020). Jurnal Transformatika	5
Εικόνα 2: Συνολική παραγωγή Bitcoin. Total Circulation of Bitcoin (2023). Blockchain.com	13
Εικόνα 3: Ιστορική τιμή του Bitcoin από το 2018. Historical Bitcoin price (2023). Blockchain.com	14
Εικόνα 4: Ιστορική τιμή του Litecoin από το 2018. Historical Litecoin price (2023). Blockchain.com	16
Εικόνα 5: Ψηφοφορία των validators για το ποιο μπλοκ θεωρούν σωστό. Validator committee (2023). ethos.dev	17
Εικόνα 6: Συνολική παραγωγή ETH. Total Ethereum Circulation (2023). Blockchain.com	17
Εικόνα 7: Ιστορική τιμή του Ethereum από το 2018. Historical Ethereum price (2023). Blockchain.com	18
Εικόνα 8: Επαληθευμένες συναλλαγές του Bitcoin ανά μέρα. Confirmed Transactions Per Day (2022). Blockchain.com	21
Εικόνα 9: Όγκος συναλλαγών Bitcoin σε USD. USD Exchange Trade Volume (2022). Blockchain.com	21
Εικόνα 10: Ανάλυση χαρακτηριστικών χρονοσειράς	26
Εικόνα 11: Μοντέλα εκθετικής εξομάλυνσης. Gardner Jr., E.S., Exponential Smoothing Models (1987). International journal of forecasting	29
Εικόνα 12: Γραφική αναπαράσταση ενός απλού μοντέλου γραμμικής παλινδρόμησης	35
Εικόνα 13: Παραδείγματα συντελεστή συσχέτισης Pearson	36
Εικόνα 14. Πίνακας εμπιστοσύνης t-values. t-Distribution Critical Values (2023).Woodstown.org.....	39
Εικόνα 15. Συντελεστές συσχέτισης ωριαίων δεδομένων των μέσων δεικτών αλγορίθμων ομοφωνίας	46
Εικόνα 16. Συντελεστές συσχέτισης ημερήσιων δεδομένων των μέσων δεικτών αλγορίθμων ομοφωνίας	47
Εικόνα 17. Ιστορικές τιμές του Bitcoin.....	48
Εικόνα 18. Ιστορικές τιμές του Cardano	48
Εικόνα 19. Ιστορικές τιμές του Ethereum	48
Εικόνα 20. Ιστορικές τιμές του Dash	48
Εικόνα 21. Τιμές κρυπτονομισμάτων ανά τον αλγόριθμο ομοφωνίας	49
Εικόνα 22. Σύγκριση του Bitcoin με άλλους χρηματιστηριακούς δείκτες	49

Εικόνα 23. Ο γενικός δείκτης PoW σε σύγκριση με τα επιμέρους κρυπτονομίσματα	51
Εικόνα 24. Ο γενικός δείκτης PoS σε σύγκριση με τα επιμέρους κρυπτονομίσματα..	51
Εικόνα 25. Ο γενικός δείκτης PoA σε σύγκριση με τα επιμέρους κρυπτονομίσματα .	52
Εικόνα 26. Ο γενικός δείκτης FBA σε σύγκριση με τα επιμέρους κρυπτονομίσματα	52
Εικόνα 27. Ο γενικός δείκτης PoW σε σχέση με το Bitcoin	53
Εικόνα 28. Τέσσερεις γενικοί δείκτες αλγορίθμου ομοφωνίας.....	55
Εικόνα 29. Πρόβλεψη του δείκτη PoW με την χρήση του μοντέλου βέλτιστης ολίσθησης	58
Εικόνα 30. Πρόβλεψη του δείκτη PoS με την χρήση του μοντέλου βέλτιστης ολίσθησης	58
Εικόνα 31. Πρόβλεψη του δείκτη FBA με την χρήση του μοντέλου βέλτιστης ολίσθησης	58
Εικόνα 32. Πρόβλεψη του δείκτη PoA με την χρήση του μοντέλου βέλτιστης ολίσθησης	58
Εικόνα 33. Πρόβλεψη του δείκτη PoS με την χρήση του μοντέλου 6.2.1	60
Εικόνα 34. Πρόβλεψη του δείκτη PoS με την χρήση του μοντέλου 6.2.2	60
Εικόνα 35. Πρόβλεψη του δείκτη FBA με την χρήση του μοντέλου 6.2.1	60
Εικόνα 36. Πρόβλεψη του δείκτη FBA με την χρήση του μοντέλου 6.2.2	60
Εικόνα 37. Πρόβλεψη του δείκτη PoA με την χρήση του μοντέλου 6.2.1	61
Εικόνα 38. Πρόβλεψη του δείκτη PoA με την χρήση του μοντέλου 6.2.2	61
Εικόνα 39. Πρόβλεψη του δείκτη PoW με την χρήση του μοντέλου 6.2.2	63
Εικόνα 40. Πρόβλεψη του δείκτη PoW με την χρήση του μοντέλου 6.2.4	63
Εικόνα 41. Πρόβλεψη του δείκτη PoS με την χρήση του μοντέλου 6.2.2	63
Εικόνα 42. Πρόβλεψη του δείκτη PoS με την χρήση του μοντέλου 6.2.4	63
Εικόνα 43. Πρόβλεψη του δείκτη PoA με την χρήση του μοντέλου 6.2.2	63
Εικόνα 44. Πρόβλεψη του δείκτη PoA με την χρήση του μοντέλου 6.2.4	63
Εικόνα 45. Πρόβλεψη του δείκτη FBA με την χρήση του μοντέλου 6.2.2	64
Εικόνα 46. Πρόβλεψη του δείκτη FBA με την χρήση του μοντέλου 6.2.4	64
Εικόνα 47. Διαφορά των μοντέλων 6.2.4 και 6.2.5 για την πρόβλεψη του PoW	65
Εικόνα 48. Πρόβλεψη του δείκτη PoW με την χρήση του μοντέλου 6.2.6	66
Εικόνα 49. Πρόβλεψη του δείκτη PoS με την χρήση του μοντέλου 6.2.6	66
Εικόνα 50. Πρόβλεψη του δείκτη FBA με την χρήση του μοντέλου 6.2.6	66
Εικόνα 51. Πρόβλεψη του δείκτη PoA με την χρήση του μοντέλου 6.2.6	66
Εικόνα 52. Πρόβλεψη του δείκτη PoW με την χρήση του μοντέλου 6.2.7	67

Εικόνα 53. Πρόβλεψη του δείκτη PoS με την χρήση του μοντέλου 6.2.7.....	67
Εικόνα 54. Πρόβλεψη του δείκτη FBA με την χρήση του μοντέλου 6.2.7	67
Εικόνα 55. Πρόβλεψη του δείκτη PoA με την χρήση του μοντέλου 6.2.7	67
Εικόνα 56. Πρόβλεψη του δείκτη PoW με την χρήση του μοντέλου 6.2.8	68
Εικόνα 57. Πρόβλεψη του δείκτη PoS με την χρήση του μοντέλου 6.2.8.....	68
Εικόνα 58. Πρόβλεψη του δείκτη FBA με την χρήση του μοντέλου 6.2.8	68
Εικόνα 59. Πρόβλεψη του δείκτη PoA με την χρήση του μοντέλου 6.2.8	68
Εικόνα 60. Κέρδος μοντέλου πρόβλεψης για όλες τις τεχνολογίες	70

Κατάλογος Πινάκων

Πίνακας 1: Τα δημοφιλέστερα ανταλλακτήρια κρυπτονομισμάτων. Νοέμβριος 2022. Coinmarketcap.com	19
Πίνακας 2: Κάποια δημοφιλή κρυπτονομίσματα. Νοέμβριος 2022. Coinmarketcap.com	19
Πίνακας 3. Παράδειγμα συντελεστών συσχέτισης των γενικών δεικτών.....	54
Πίνακας 4. Βέλτιστα διαστήματα ολισθήσης.....	57
Πίνακας 5. Αποτελέσματα δεικτών ακρίβειας μοντέλου βέλτιστων ολισθήσεων	58
Πίνακας 6. Αποτελέσματα δεικτών ακρίβειας μοντέλου βέλτιστων ολισθήσεων και ολισθήσεων κατά μία ημέρα.....	59
Πίνακας 7. Αποτελέσματα δεικτών ακρίβειας μοντέλου βέλτιστων ολισθήσεων, ολισθήσεων κατά μία ημέρα και cherry-picked ημερών.....	61
Πίνακας 8. Αποτελέσματα δεικτών ακρίβειας μοντέλου επανεκπαίδευσης	62
Πίνακας 9. Αποτελέσματα δεικτών ακρίβειας μοντέλου επανεκπαίδευσης 6.2.5	64
Πίνακας 10. Αποτελέσματα δεικτών ακρίβειας μοντέλου με προσθήκη του S&P500	66
Πίνακας 11. Αποτελέσματα δεικτών ακρίβειας μοντέλου με προσθήκη του αργού πετρελαίου	68
Πίνακας 12. Αποτελέσματα δεικτών ακρίβειας μοντέλου με προσθήκη του χρυσού	69
Πίνακας 13. Σύγκριση αποτελεσμάτων επένδυσης του βέλτιστου μοντέλου	70
Πίνακας 14. Σύγκριση αποτελεσμάτων επένδυσης σε επιμέρους κρυπτονομίσματα..	71

Κεφάλαιο 1. Εισαγωγή

1.1 Αντικείμενο εργασίας

Μετά την ραγδαία άνοδο και την ύστερη πτώση των κρυπτονομισμάτων τα τελευταία χρόνια, το Bitcoin και τα υπόλοιπα κρυπτονομίσματα έχουν αρχίσει να θεωρούνται περισσότερο επενδυτικά αγαθά παρά εναλλακτικά νομίσματα. Ο δελεαστικότερος παράγοντας για επένδυση σε αυτά ήταν η τεράστια αύξηση της τιμής τους κατά την διάρκεια που έγιναν δημοφιλή προσφέροντας απίστευτα κέρδη στους πρώτους επενδυτές. Παρόλα αυτά λόγω της ασταθούς φύσης τους αυξάνεται η σημασία μίας καλής πρόβλεψης, για την άντληση ορθών αποφάσεων επένδυσης. Τα τελευταία χρόνια έχουν αναπτυχθεί πολλές τεχνικές πρόβλεψης αυτών με κυριότερες την μηχανική μάθηση. Παρόλα αυτά οι περισσότερες έρευνες ασχολήθηκαν με το δημοφιλέστερο κρυπτόνμισμα – το Bitcoin – και με την πρόβλεψη της τιμής του ίδιου, βάσει των παρελθοντικών τιμών του.

Ένα πολύ μικρό κομμάτι των ερευνών ασχολήθηκε με την σχέση μεταξύ διαφορετικών κρυπτονομισμάτων, και πιο ειδικά με την σύγκριση των αλγορίθμων ομοφωνίας που αυτά χρησιμοποιούν, χωρίς όμως εκτενή σύγκριση των δημοφιλέστερων αυτών τεχνολογιών. Η αξία που έχει μία μετοχή στο χρηματιστήριο προέρχεται από την αξία του προϊόντος ή της εταιρίας που αυτή αντιπροσωπεύει καθώς και της μελλοντικής προοπτικής της. Εγγενώς, η αξία των κρυπτονομισμάτων – αν εξαιρεθεί η επένδυση σε αυτά για βραχυπρόθεσμο κέρδος – είναι η χρήση τους σαν συμβατικά νομίσματα. Ο μεγαλύτερος περιορισμός όμως είναι η ταχύτητα επαλήθευσης μίας συναλλαγής, λόγω των αλγορίθμων ομοφωνίας που αυτά χρησιμοποιούν. Έτσι τα πιο χρήσιμα νομίσματα και αυτά με την μεγαλύτερη «αξία» θα έπρεπε να είναι τα πιο αποδοτικά με την μεγαλύτερη ευκολία χρήσης.

Στην συγκεκριμένη διπλωματική εργασία, λοιπόν, εξετάζεται η σχέση της αξίας των κρυπτονομισμάτων βάσει της κατηγορίας αλγορίθμου ομοφωνίας (consensus) που αυτά χρησιμοποιούν για την πρόβλεψη της τιμής τους. Από της πιο σημαντικές ερωτήσεις που προσπαθούμε να απαντήσουμε είναι το αν διαφορετικοί τύποι νομισμάτων έχουν διαφορετική χρηματιστηριακή συμπεριφορά. Πιο συγκεκριμένα, οι τελικές προβλέψεις για κάποια τεχνολογία κρυπτονομισμάτων βασίζεται σε διαφορετικές τεχνολογίες κρυπτονομισμάτων με σκοπό την άντληση δεδομένων για την μεταξύ τους σχέση.

Επιπρόσθετα, η ομαδοποίηση και πρόβλεψη της τιμής βάσει κάποιων κοινών χαρακτηριστικών προσπαθεί να εξαλείψει τον παράγοντα της αστάθειας που υπάρχει στα κρυπτονομίσματα. Παράλληλα προσπαθούν να αντληθούν και οι παράγοντες που επηρεάζουν την τιμή των κρυπτονομισμάτων, καθώς αποτελούν διαφορετικό επενδυτικό αγαθό από τα συμβατικά αγαθά που έχουν ήδη μελετηθεί. Σε αυτό το σημείο χρησιμοποιούνται κάποια κοινά αγαθά (π.χ. πετρέλαιο, χρυσός), που βιβλιογραφικά φαίνεται να έχουν έστω και έναν μικρό βαθμό συσχέτισης με τα κρυπτονομίσματα, για την βελτίωση του μοντέλου πρόβλεψης και μελετώνται οι επιπτώσεις σε αυτό.

1.2 Οργάνωση εργασίας

Στο δεύτερο κεφάλαιο της διπλωματικής εργασίας παρουσιάζονται οι σημαντικότερες έννοιες των συναλλαγών και των κρυπτονομισμάτων καθώς και των βαθύτερων τεχνολογιών που δομούν τα κρυπτονομίσματα με κυριότερη από αυτές το Blockchain. Γίνεται λοιπόν εισαγωγή με κάποιες βασικές έννοιες και μετά εξετάζεται η δομή του Blockchain και των επιμέρους στοιχείων του, που συνθέτουν τα θεμέλια των κρυπτονομισμάτων. Έπειτα ορίζονται και επιπρόσθετοι όροι των κρυπτονομισμάτων όπως η εξόρυξη και το πορτοφόλι, για να γίνει ευρύτερα κατανοητός ο τρόπος διάρθρωσης τους. Τέλος αναφέρονται και οι πιο γνωστοί αλγόριθμοι ομοφωνίας για την επαλήθευση των επιμέρους συναλλαγών.

Το τρίτο κεφάλαιο ξεκινάει με μία ιστορική αναδρομή πάνω στα κρυπτονομίσματα, αναφέροντας λεπτομέρειες για την δημιουργία τους και κάποια σημαντικά ορόσημα, για τα δημοφιλέστερα αυτών. Επίσης παρουσιάζεται η εικόνα της σημερινής αγοράς με ενδεικτικές τιμές για την κύμανση της τιμής των κρυπτονομισμάτων τα τελευταία χρόνια. Παράλληλα, γίνεται και μία μικρή αναφορά στο χρηματιστήριο και ανταλλακτήρια κρυπτονομισμάτων, τονίζοντας την χρήση των κρυπτονομισμάτων ως επενδυτικό αγαθό. Τέλος γίνεται μία μικρή βιβλιογραφική ανασκόπηση της συνάφειας των κρυπτονομισμάτων μεταξύ τους και με άλλα επενδυτικά αγαθά.

Το τέταρτο κεφάλαιο αναφέρεται στις πιο συχνές τεχνικές προβλέψεων που χρησιμοποιούνται σήμερα. Γίνεται αναφορά τόσο στις ποσοτικές όσο και στις ποιοτικές μεθόδους πρόβλεψης, επικεντρώνοντας το ενδιαφέρον κυρίως στην πρώτη κατηγορία, και πιο συγκεκριμένα στις μεθόδους χρονοσειρών. Σε αυτό το κομμάτι γίνεται εκτενής αναφορά στα μοντέλα γραμμικής παλινδρόμησης που θα χρησιμοποιηθούν και στο πειραματικό μέρος της διπλωματικής εργασίας. Σημαντικό μέρος των μοντέλων πρόβλεψης αποτελούν και οι μέθοδοι αξιολόγησης των μοντέλων αυτών, και εδώ αναφέρονται και οι δείκτες ακρίβειας, που είναι μαθηματικές σχέσεις που προσπαθούν να χαρακτηρίσουν την ακρίβεια της πρόβλεψης του κάθε μοντέλου. Γίνεται αναφορά στις κύριες διαφορές μεταξύ τους και στην περίπτωση που προτιμάται ο κάθε δείκτης.

Στο πέμπτο κεφάλαιο παρουσιάζεται η πειραματική διαδικασία. Αρχικά γίνεται αναφορά στον τρόπο συγκέντρωσης των δεδομένων, από το Yahoo Finance, και στην προ επεξεργασία αυτών. Έπειτα γίνεται μικρή παρουσίαση των αρχικών δεδομένων καθώς και μία υποτυπώδης σύγκριση για να φανεί η γραμμική συσχέτιση που υπάρχει σε αυτά. Στο επόμενο στάδιο παρουσιάζεται ο τρόπος δημιουργίας μίας μέσης μετρικής για κάθε τεχνολογία (βάσει του αλγορίθμου ομοφωνίας του κάθε κρυπτονομίσματος) και συγκρίνεται γραφικά η γραμμική συσχέτιση των μετρικών, σε σχέση με τις πρωτογενείς τιμές. Στο τέλος υπολογίζεται η προαναφερθείσα γραμμική συσχέτιση για την εύρεση του καλύτερου προβλεπτικού διαστήματος για το μοντέλο που θα χρησιμοποιηθεί στο επόμενο κεφάλαιο.

Στο έκτο κεφάλαιο παρουσιάζεται ο τρόπος χρήσης των δεδομένων που αναφέρθηκαν στο πέμπτο κεφάλαιο για την δημιουργία των προβλέψεων. Τα δεδομένα αυτά τροφοδοτούνται σε ένα μοντέλο απλής γραμμικής παλινδρόμησης, στο οποίο προστίθενται σε στάδια όλο και περισσότερα δεδομένα. Τα αποτελέσματα στα διάφορα στάδια συγκρίνονται μεταξύ τους, με σκοπό την άντληση κάποιων γενικών συμπερασμάτων για τις σχέσεις των δεδομένων καθώς και για την εύρεση του

υποσυνόλου δεδομένων που δίνει τα καλύτερα αποτελέσματα. Τέλος αξιολογούνται τα αποτελέσματα του καλύτερου μοντέλου και αναφέρονται τα συμπεράσματα που προκύπτουν.

Στο έβδομο κεφάλαιο παρουσιάζονται τα συμπεράσματα που προέκυψαν καθ' όλη την διάρκεια της μελέτης. Παράλληλα προτείνονται και κάποιες μελλοντικές προεκτάσεις για πιθανές βελτιώσεις που μπορεί να δεχτεί το μοντέλο πρόβλεψης που παρουσιάζεται.

Κεφάλαιο 2. Μελέτη των κρυπτονομισμάτων

2.1 Βασικές Έννοιες

1. Εικονικό νόμισμα

Με τον όρο εικονικό νόμισμα ορίζουμε έναν ψηφιακό συνάλλαγμα που δεν υπάγεται σε κανονισμούς, εκδίδεται και ελέγχεται από τον δημιουργό του. Τέτοια νομίσματα αποτελούν μορφή συναλλάγματος για τα μέλη της ψηφιακής κοινότητας που τα χρησιμοποιούν (European Central Bank, 2018).

2. Ηλεκτρονικό χρήμα

Με τον όρο ηλεκτρονικό χρήμα περιγράφουμε κάθε μορφή μεταφοράς κεφαλαίου μεταξύ δύο ή περισσότερων μερών που γίνεται με ψηφιακό τρόπο και χωρίς τη μεσολάβηση υλικού μέσου.

3. Κρυπτονόμισμα

Τα κρυπτονομίσματα είναι ψηφιακά νομίσματα τα οποία βασίζονται σε αποκεντρωμένο λογισμικό και χρησιμοποιούν την τεχνολογία του Blockchain.

4. Ανταλλακτήριο κρυπτονομισμάτων

Τα ανταλλακτήρια κρυπτονομισμάτων αποτελούν ιστοσελίδες που επιτρέπουν την ανταλλαγή κρυπτονομισμάτων με άλλα ψηφιακά αγαθά ή συναλλάγματα.

5. Χρηματιστηριακή αξία (Market Capitalization)

Η χρηματιστηριακή αξία αποτελεί την τιμή μίας μετοχής πολλαπλασιασμένη με τον αριθμό των συνολικών μετοχών που είναι διαθέσιμες στην δημόσια αγορά. Από την στιγμή που οι μετοχές ανταλλάσσονται σε δημόσιες αγορές, η χρηματιστηριακή δηλώνει κατά κάποιο τρόπο την κοινή γνώμη για την αξία μίας εταιρίας.

2.2 Η έννοια του Blockchain

Το blockchain είναι μία κατακευματισμένη βάση αντιγράφων όλων των ψηφιακών συναλλαγών που έχουν γίνει μεταξύ των συμμετεχόντων μελών στο σύστημα. Κάθε συναλλαγή στο πρέπει να επαληθευθεί από την πλειοψηφία των συμμετεχόντων για να εισαχθεί στο blockchain (Crosby, et al., 2016). Η συγκεκριμένη τεχνολογία παρέχει στους χρήστες την δυνατότητα να πιστοποιούν συναλλαγές με απουσία κεντρικού ελέγχου. Παρακάτω παρουσιάζεται η βασική δομή του Blockchain.

1. Hash function

Hash function ονομάζεται μία συνάρτηση που κωδικοποιεί δεδομένα σε μία ακολουθία σταθερού μήκους. Το αποτέλεσμα της συνάρτησης λέγεται hash.

2. Λογιστικό βιβλίο (Block)

Τα block είναι δομές που περιλαμβάνουν επικυρωμένες συναλλαγές οι οποίες έχουν περάσει από hash functions και αποθηκεύονται σε δομές δεδομένων που ονομάζονται δέντρα Merkle (The Economist, 2015). Κάθε block περιλαμβάνει το κρυπτογραφημένο hash του προηγούμενου του block, που συνδέει τα block μεταξύ τους, δημιουργώντας την αλυσίδα του Blockchain. Το πρώτο block του Blockchain που συνδέει την αλυσίδα

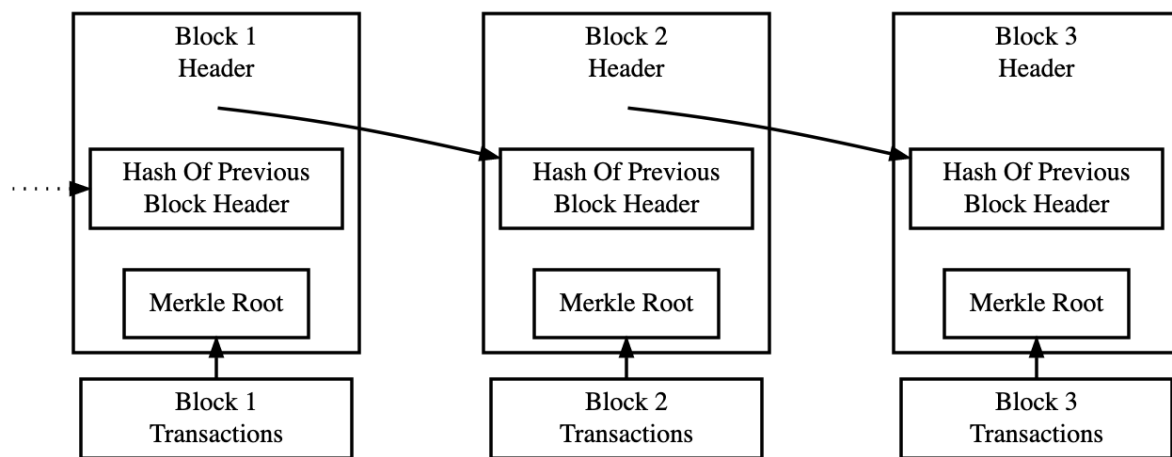
ονομάζεται genesis block (Bhaskar & Chuen, 2015). Το κάθε block περιέχει δεδομένα για πολλές συναλλαγές που έχουν γίνει στο δίκτυο. Τα blocks έχουν πεπερασμένη χωρητικότητα η οποία εξαρτάται από το εκάστοτε κρυπτονόμισμα και άρα περιέχουν διαφορετικό πλήθος συναλλαγών.

3. Κόμβος (Node)

Ως κόμβος ορίζεται μία υπολογιστική μονάδα στο blockchain. Κάθε κόμβος είναι υπεύθυνος για την πιστοποίηση των συναλλαγών, την διατήρηση και επικαιροποίηση της αλυσίδας.

4. Αλυσίδα μπλοκ (Blockchain)

Ορίζεται ως η αλυσίδα των blocks, διατεταγμένα σε χρονολογική σειρά, και πρακτικά είναι το μητρώο όλων των συναλλαγών του συστήματος.



Simplified Bitcoin Block Chain

Εικόνα 1: Απεικόνιση ιεραρχίας αλυσίδας block. Dzulfikar, F. & Susanto, A., *Simplified Bitcoin Block Chain* (2020). *Jurnal Transformatika*

Το Blockchain χωρίζεται σε 3 βασικές κατηγορίες:

a. Δημόσιο Blockchain

Στο δημόσιο Blockchain ο κάθε χρήστης έχει πρόσβαση και δυνατότητα δημιουργίας κόμβων καθώς και πραγματοποίησης συναλλαγών και επαλήθευσης τους. Το δημόσιο Blockchain αποτελεί την πιο διαδεδομένη μορφή του Blockchain καθώς αποτελεί την βασική δομή στην οποία βασίζεται η τεχνολογία των κρυπτονομισμάτων. Το πλεονέκτημα του δημόσιου Blockchain είναι ότι δεν ελέγχεται από μία μερίδα κόμβων, αλλά από όλους τους χρήστες. Παρόλα αυτά ένα τέτοιο σύστημα χρειάζεται μεγάλη ποσότητα ενέργειας για την επαλήθευση των συναλλαγών, ένα από τα κυριότερα μειονεκτήματα της τεχνολογίας του Blockchain.

b. Ιδιωτικό Blockchain

Το ιδιωτικό Blockchain αποτελεί ένα κλειστό σύστημα όπου μία μεμονωμένη αρχή ελέγχει τους κανόνες του Blockchain. Χαρακτηριστικό παράδειγμα ιδιωτικού Blockchain αποτελεί η πλατφόρμα “Multichain” η οποία παρέχει τη δυνατότητα σε

εταιρείες, κυρίως πολυεθνικές, να φτιάξουν ιδιωτικά Blockchain, με σκοπό την βελτίωση των συναλλαγών στο εσωτερικό τους. Το ιδιωτικό Blockchain παρόλα αυτά δεν έχει μεγάλη χρησιμότητα.

c. Ομοσπονδιακό Blockchain

Σε αντίθεση με το δημόσιο και το ιδιωτικό Blockchain, εδώ, υπάρχει μία ομάδα χρηστών που παίρνουν τις αποφάσεις και έχουν τα δικαιώματα του δικτύου – σε αντίθεση με το ιδιωτικό, όπου ο χρήστης με τα δικαιώματα είναι μοναδικός. Οι χρήστες αυτοί επιλέγονται από πριν και νέοι χρήστες δεν μπορούν να εισέλθουν στο σύστημα. Στην συγκεκριμένη υλοποίηση ο κάθε κόμβος έχει την ίδια δύναμη για την επαλήθευση συναλλαγών. Η πιο διάσημη υλοποίηση ομοσπονδιακού Blockchain αποτελεί ο Hyperledger, μία πλατφόρμα που προσφέρει εφαρμογές Blockchain σε εταιρίες. Από την στιγμή που οι συμμετέχοντες κόμβοι είναι προκαθορισμένοι, το ομοσπονδιακό Blockchain δεν επικεντρώνεται σε χρηματικές συναλλαγές, καθώς δεν μπορεί να αποκτήσει νέους χρήστες. Το κύριο πλεονέκτημα του σε σχέση με το δημόσιο Blockchain είναι η απόδοση και η ταχύτητα του, που οδηγούν σε λιγότερη κατανάλωση ενέργειας.

5. Διαδικασία Εισαγωγής στο Blockchain

Πολλές συναλλαγές συγκεντρώνονται μαζί και δημιουργούν ένα block. Αυτό το block κρυπτογραφείται, κωδικοποιείται και διαμοιράζεται σε όλους τους κόμβους του δικτύου. Όταν επιβεβαιωθεί από τους χρήστες ότι το block περιέχει τις σωστές συναλλαγές εισάγεται μετά το τελευταίο block και συνεχίζεται η αλυσίδα.

Σήμερα η τεχνολογία του Blockchain έχει πολλές εφαρμογές, αλλά η σημαντικότερη από αυτές είναι η χρήση του σε πολλά κρυπτονομίσματα.

2.3 Εφαρμογές του Blockchain

1. Οικονομία

Η τεχνολογία του Blockchain έγινε δημοφιλής με την χρήση της για την υλοποίηση των πρώτων κρυπτονομισμάτων. Σε αυτή την μορφή, το πλεονέκτημα τους είναι η δημιουργία ενός εναλλακτικού νομίσματος το οποίο δεν ελέγχεται από μία μεμονωμένη αρχή και παρέχει διαφάνεια και ασφάλεια στους χρήστες.

2. Ασφάλιση

Λόγω της ασφαλούς φύσης του, το Blockchain μπορεί να χρησιμοποιηθεί για την επαλήθευση δεδομένων σε ασφαλιστικά συμβόλαια καθώς και την αποτροπή ξεπλύματος χρήματος. Ο σημαντικότερος παράγοντας σε αυτές τις εφαρμογές είναι τα έξυπνα συμβόλαια, ένας τρόπος αυτόματης τήρησης των όρων ενός συμβολαίου.

3. Κανονιστική συμμόρφωση και Έλεγχος (Regulatory compliance & Audit)

Η δομή του Blockchain προσφέρει ασφάλεια, αποφυγή ανθρώπινων σφαλμάτων και διαφάνεια των οικονομικών αρχείων σε εφαρμογές λογιστικής. Το μεγαλύτερο πλεονέκτημα σε αυτές τις εφαρμογές είναι ότι κανένας δεν μπορεί να αλλάξει τα αρχεία από την στιγμή που θα προστεθούν στο σύστημα.

4. Υγεία

Τα στοιχεία του ασθενή μπορούν να αποθηκευτούν ασφαλώς σε τέτοια συστήματα παρέχοντας την δυνατότητα άμεσης πρόσβασης δεδομένων σε πολλές ιατρικές εφαρμογές.

5. Αγορά ακινήτων

Ορισμένα από τα θέματα αγοράς και πώλησης ακινήτων είναι η γραφειοκρατία, η έλλειψη διαφάνειας, η απάτη και τα λάθη στα δημόσια αρχεία. Επομένως, η χρήση τεχνολογίας Blockchain μπορεί να βοηθήσει στην παρακολούθηση, στην επαλήθευση της ιδιοκτησίας, στην εξασφάλιση της ακρίβειας των εγγράφων και στη μεταβίβαση των πράξεων ιδιοκτησίας.

2.4 Τεχνολογία των κρυπτονομισμάτων

1. Προσωπικό & Δημόσιο Κλειδί (public & private key)

Τα κλειδιά αυτά χρησιμοποιούνται μία μορφή κρυπτογράφησης που ονομάζεται κρυπτογραφία προσωπικού κλειδιού (public-key cryptography). Το ζεύγος δημιουργείται μέσα από κρυπτογραφικούς αλγορίθμους. Το δημόσιο κλειδί είναι γνωστό σε όλους τους χρήστες και μπορεί να χρησιμοποιηθεί για την κωδικοποίηση ενός μηνύματος, ενώ το προσωπικό κλειδί ανήκει μόνο στον ίδιο τον χρήστη και χρησιμοποιείται για την αποκρυπτογράφηση των μηνυμάτων. (Menezes, et al., 1996) (Bernstein, 2008). Στα κρυπτονομίσματα το δημόσιο κλειδί χρησιμοποιείται για να λάβει ο χρήστης κρυπτονομίσματα, ενώ το προσωπικό κλειδί χρησιμοποιείται για να στείλει.

2. Πορτοφόλι κρυπτονομισμάτων (Cryptocurrency wallet)

Ονομάζεται το μέσο αποθήκευσης των διευθύνσεων του ζεύγους κλειδιών (private και public) που χρησιμοποιούνται για να σταλούν ή να ληφθούν κρυπτονομίσματα.

3. Εξόρυξη (Mining)

Για την ασφάλεια του συστήματος, η επαλήθευση των συναλλαγών στα κρυπτονομίσματα, αποτελεί μία εργασία που απαιτεί τεράστια υπολογιστική ισχύ. Οι χρήστες που κάνουν την επαλήθευση αυτή, ως αντάλλαγμα λαμβάνουν κρυπτονομίσματα. Η διαδικασία αυτή ονομάζεται εξόρυξη. Κατά την διαδικασία της εξόρυξης, το υπολογιστικό σύστημα καλείται να εκτελέσει πολύπλοκους μαθηματικούς υπολογισμούς και στο τέλος, ο κόμβος που θα βρει το αποτέλεσμα ανταμείβεται με κάποιο αριθμό κρυπτονομισμάτων.

4. Ομότιμη σύνδεση (Peer-to-peer - P2P)

Η ομότιμη σύνδεση ή P2P αναφέρεται στην επικοινωνία ανάμεσα σε δύο ή περισσότερα

μέλη ενός έντονα διασυνδεδεμένου δικτύου. Οι συμμετέχοντες ενός τέτοιου δικτύου επικοινωνούν απευθείας μεταξύ τους, μέσα από ένα μοναδικό σημείο διαμεσολάβησης. Όλοι οι κόμβοι του δικτύου έχουν ίσα δικαιώματα. Πληροφορίες που βρίσκονται στον ένα κόμβο, ανάλογα με τα δικαιώματα που καθορίζονται, μπορούν να διαβαστούν από όλους τους άλλους και αντίστροφα.

5. Έξυπνα συμβόλαια (Smart contracts)

Με τον όρο έξυπνο συμβόλαιο ονομάζουμε ένα πρόγραμμα που θα εκτελεστεί αυτόματα, μετά την κάλυψη των όρων στους οποίους έχει δημιουργηθεί. Έτσι τα συμβόλαια γράφονται σε μορφή κώδικα στο Blockchain και εκτελούνται όταν έχουν καλυφθεί οι προϋποθέσεις τους. Τα έξυπνα συμβόλαια εξυπηρετούν στην συναλλαγή χρημάτων, ακινήτων, μετοχών και οποιουδήποτε αγαθού ή υπηρεσίας με διαφανή τρόπο, αποφεύγοντας παράλληλα τις υπηρεσίες κάποιου μεσάζοντα. Πιο συγκεκριμένα, τα έξυπνα συμβόλαια όχι μόνο καθορίζουν τους κανόνες και τις κυρώσεις γύρω από μια συμφωνία με τον ίδιο τρόπο που εφαρμόζει μια παραδοσιακή σύμβαση, αλλά και επιβάλλουν αυτομάτως αυτές τις υποχρεώσεις και έτσι προσφέρουν την δυνατότητα αυτόματης προσαρμογής, ανάλογα με τους όρους και τις συνθήκες της αγοράς.

6. Αποκεντρωμένη ομοφωνία (Decentralized consensus)

Εξαιτίας της αποκεντρωμένης φύσης του Blockchain, κόμβοι που δεν είναι απαραίτητα έμπιστοι, χρησιμοποιούνται για την επαλήθευση των συναλλαγών. Στο Blockchain, η ομοφωνία πρακτικά είναι η ανάθεση της επαλήθευσης των συναλλαγών σε τοπικούς κόμβους, που πρέπει να επιβεβαιώσουν ότι η συναλλαγή πληροί κάποιους κανόνες για να εισαχθεί στην αλυσίδα. Έτσι δεν υπάρχει κάποια αρμόδια αρχή που διέπει και ελέγχει τους κόμβους και θέτει ορισμένους κανόνες που πρέπει να τηρούνται. Συνεπώς δεν υπάρχει μία πηγή αλήθειας στο σύστημα, αλλά μία συναλλαγή θεωρείται έγκυρη αν επαληθευτεί από την πλειοψηφία των κόμβων του συστήματος. (Tasca & Tessone, 2017).

7. Ομοφωνία (Consensus)

Όπως αναφέρθηκε και προηγουμένως, η ομοφωνία αναφέρεται στους κανόνες που διέπουν το Blockchain, και χρησιμοποιούνται για την ανανέωση του με νέες συναλλαγές με στόχο την εγκυρότητα των δεδομένων του. Υπάρχουν διαφορετικές μέθοδοι ομοφωνίας αναλόγως με την δομή του κάθε κρυπτονομίσματος.

2.5 Βασικές κατηγορίες κρυπτονομισμάτων βάσει αλγορίθμου ομοφωνίας

Τα κρυπτονομίσματα χωρίζονται σε κατηγορίες ανάλογα με τους αλγορίθμους ομοφωνίας που χρησιμοποιούν, για την επαλήθευση των συναλλαγών. Οι κυριότεροι αυτών αναφέρονται παρακάτω.

2.5.1 Απόδειξη εργασίας (Proof-of-Work)

Ο αλγόριθμος Proof-of-Work (PoW) αποτελεί τον τρόπο επιβεβαίωσης ότι ένας κόμβος έχει πραγματοποιήσει μια συγκεκριμένη εργασία που χρειάστηκε σημαντικό ποσό υπολογιστικών πόρων. (Jakobsson & Juels, 1999) Ο αλγόριθμος έχει χρησιμοποιηθεί σαν πρωτόκολλο ασφαλείας σε πολλές εφαρμογές. Μία από αυτές είναι και κάποια κρυπτονομίσματα που τον χρησιμοποιούν για την επαλήθευση των συναλλαγών πριν την εγγραφή τους στο Blockchain. Το πιο διαδεδομένο κρυπτονόμισμα, το Bitcoin, χρησιμοποιεί τον αλγόριθμο για την επαλήθευση των συναλλαγών του. Στο σύστημα αυτό, οι κόμβοι, οι οποίοι λέγονται miners, συνδέονται μεταξύ τους σε ένα δίκτυο και επαληθεύουν τις συναλλαγές. Για την εισαγωγή του block στο Blockchain, ο miner θα πρέπει να αποκρυπτογραφήσει το hash του block, δοκιμάζοντας πολλούς πιθανούς συνδυασμούς. Στο PoW η πιθανότητα ένας miner να κάνει εξόρυξη (mine) του block

εξαρτάται από την υπολογιστική ισχύ που θα διαθέσει ο ίδιος καθώς και την συνολική υπολογιστική ισχύ που διατίθεται από τους υπόλοιπους κόμβους του δικτύου. Πιο συγκεκριμένα οι miners πρέπει να βρουν την λύση μίας hash function, δοκιμάζοντας hash τιμές που υπολογίζουν βάσει του συνδυασμού των προηγούμενων τιμών hash, των συναλλαγών που βρίσκονται στο block και μία τυχαία ακολουθία που αλλάζουν συνεχώς (nonce). Η λύση που θα βρεθεί θα πρέπει, είναι τέτοια ώστε η καινούργια τιμή hash που θα υπολογιστεί να ξεκινάει από ένα συγκεκριμένο αριθμό μηδενικών (Tasca & Tessone, 2017). Από αυτά, είναι κατανοητό πόσο μεγάλη υπολογιστική ισχύ και ενέργεια ξοδεύεται για την επαλήθευση κάθε block συναλλαγών.

2.5.2 Απόδειξη κυριότητας (Proof-of-Stake)

Ο αλγόριθμος απόδειξης κυριότητας (PoS) προσπαθεί να λύσει το πρόβλημα δαπάνης ενέργειας που έχει ο αλγόριθμος PoW. Για την επαλήθευση της συναλλαγής, ο αλγόριθμος κυριότητας επιλέγει τυχαία έναν κόμβο του συστήματος για να επαληθεύσει την συναλλαγή. Ο τρόπος επιλογής του κόμβου συσχετίζεται με το πλήθος των ψηφιακών αγαθών (κρυπτονομισμάτων) που διαθέτει ο επαληθευτής στο σύστημα. Αναλόγως το κρυπτονόμισμα, υπάρχουν διαφορετικές εφαρμογές του αλγορίθμου. Μία εφαρμογή του αλγορίθμου που έχει υλοποιηθεί σε κάποια κρυπτονομίσματα (NXT και BlackCoin), συνυπολογίζει το μικρότερο hash μαζί με το πλήθος των κρυπτονομισμάτων που κατέχει ο κόμβος για να βρει τον επαληθευτή. Ένας άλλος τρόπος επαλήθευσης, υλοποιημένος στο Peercoin, είναι η χρήση του νοήματος της «ηλικίας» των νομισμάτων. Στο συγκεκριμένο σύστημα ο κόμβος επιλέγεται μέσω του υπολογισμού του γινομένου του πλήθους των νομισμάτων του κόμβου και του χρονικού διαστήματος που ο κόμβος κατέχει αυτά τα νομίσματα. Η λογική του αλγορίθμου αυτού, είναι ότι όσα περισσότερα ψηφιακά αγαθά κατέχει ο κόμβος στο σύστημα, τόσο πιο έμπιστος είναι. Παρόλο που ο αλγόριθμος λύνει το πρόβλημα της κατανάλωσης της ενέργειας, δημιουργείται ένα καινούργιο πρόβλημα, της εμπιστοσύνης στον κόμβο που είναι υπεύθυνος για την επαλήθευση των συναλλαγών. Το πρόβλημα αυτό λύνεται με τον αλγόριθμο Delegated Proof-of-Stake (DPoS). Σε αυτή την μορφή του PoS εκλέγονται αρκετές ομάδες κόμβων για την δημιουργία και επαλήθευση των μπλοκ, που συνεργάζονται μεταξύ τους. Αυτοί οι επιλεγμένοι κόμβοι ονομάζονται μάρτυρες (witnesses). Οι μάρτυρες ελέγχουν ο ένας τον άλλο και υπάρχουν εναλλακτικοί κόμβοι σε περίπτωση κάποιου σφάλματος ή κακοήθης συμπεριφοράς κάποιου επιλεγμένου κόμβου.

2.5.3 Απόδειξη Εξουσίας (Proof-of-Authority)

Ο αλγόριθμος εξουσίας (PoA) χρησιμοποιεί συγκεκριμένους προκαθορισμένους κόμβους για την επαλήθευση των κόμβων. Ο αλγόριθμος χρησιμοποιείται κυρίως σε ιδιωτικά δίκτυα όπου οι κόμβοι επαλήθευσης είναι προκαθορισμένες πραγματικές οντότητες, εκλεγμένες από τους δημιουργούς του Blockchain.

2.5.4 Απόδειξη Χωρητικότητας (Proof-of-Capacity)

Ο αλγόριθμος χωρητικότητας (PoC) χρησιμοποιεί την χωρητικότητα των υπολογιστικών συστημάτων σαν πόρο του δικτύου. Σε αυτή την υλοποίηση δεν δίνεται έμφαση στην υπολογιστική ισχύ του επεξεργαστή, αλλά στον αποθηκευτικό χώρο που διαθέτει το σύστημα. Οι κόμβοι στην διαδικασία της εξόρυξης δεσμεύουν χώρο στον

δίσκο τους, όπου όσο μεγαλύτερος χώρος δεσμευτεί, τόσο μεγαλύτερη είναι και η πιθανότητα εξόρυξης του block από τον κόμβο. Το θετικό του αλγορίθμου αυτού είναι ότι μειώνεται η υπολογιστική ισχύ που ξοδεύεται στην διαδικασία επαλήθευσης και έτσι ο αλγόριθμος είναι περισσότερο αποδοτικός όσον αφορά την ενέργεια σε σχέση με τον PoW.

2.5.5 Υβριδικοί Αλγόριθμοι (Hybrid Algorithms)

Υπάρχουν και κρυπτονομίσματα που χρησιμοποιούν κάποια σύνθεση δύο η περισσότερων από τους προαναφερθέντες αλγορίθμους για την επαλήθευση των block. Για παράδειγμα το Peercoin χρησιμοποιεί ένα υβριδικό μοντέλο με τους PoW και PoS.

2.5.6 Αλγόριθμος αντοχής σε βυζαντινά σφάλματα (Byzantine Fault Tolerance - BFT)

Οι αλγόριθμοι αντοχής σε βυζαντινά σφάλματα είναι αλγόριθμοι που αποτρέπουν τα σφάλματα που μπορεί να προκύψουν σε ένα καταναμημένο σύστημα από κακόβουλους κόμβους και λάθη του ίδιου του δικτύου. Μια κατηγορία αλγορίθμων βυζαντινών σφαλμάτων είναι το Federated Byzantine Agreement (FBA) που επιτρέπει την ελεύθερη εισαγωγή κόμβων στο σύστημα. Τέτοιοι αλγόριθμοι προσφέρουν μεγαλύτερη απόδοση, ευκολότερη επεκτασιμότητα και μικρά κόστη συναλλαγής. Στα συστήματα FBA οι κόμβοι επαλήθευσης διαλέγουν ποιους άλλους κόμβους εμπιστεύονται και σχηματίζουν μια ομάδα κόμβων. Το σύνολο αυτών των ομάδων επαληθεύει όλες τις συναλλαγές στο σύστημα.

Κεφάλαιο 3. Ιστορική αναδρομή και βιβλιογραφική ανασκόπηση

3.1 Ιστορική αναδρομή

Το πρώτο κρυπτονόμισμα προτάθηκε το 2008 από τον Satoshi Nakamoto, ο οποίος δημοσίευσε το άρθρο του με τίτλο “Bitcoin: A Peer-To-Peer Electronic Cash System”. Το άρθρο παρουσίαζε έναν τύπο εικονικού νομίσματος αποκεντρωμένων ψηφιακών συναλλαγών χωρίς την χρήση κάποιου χρηματοπιστωτικού ιδρύματος (Crosby, et al., 2016). Το κύριο πρόβλημα που προσπαθούσε να λύσει η δομή του Bitcoin, ήταν η έμφυτη αδυναμία του μοντέλου της εμπιστοσύνης σε έναν ενδιάμεσο προμηθευτή για την διεκπεραίωση της οποιασδήποτε συναλλαγής. Στις 9 Νοεμβρίου του 2008 το project κατοχυρώθηκε στο sourceforge.net και η πρώτη συναλλαγή έγινε στις 12 Ιανουαρίου του 2009. Τον Αύγουστο του 2010 ανακαλύφθηκε ένα σφάλμα του συστήματος που κακόβουλοι χρήστες μπόρεσαν να εκμεταλλευτούν για να κλέψουν περισσότερα από 184δισ bitcoins. Το τρωτό σημείο του συστήματος, ήταν η επαλήθευση των συναλλαγών πριν την είσοδο τους στο σύστημα. Το σφάλμα διορθώθηκε σε λίγες ώρες, και οι κακόβουλες συναλλαγές διαγράφηκαν. Αυτή ήταν το μόνο σφάλμα στην ιστορία του Bitcoin που είχε ως αποτέλεσμα εκμετάλλευση από κακόβουλους χρήστες σε τέτοιο βαθμό. Η πρώτη μεγάλη συναλλαγή που πραγματοποιήθηκε ήταν τον Μάιο του 2010, όπου αγοράστηκαν 2 πίτσες για 10000 bitcoins. Παράλληλα το 2010 ιδρύθηκε η εταιρία Mt. Gox, και το 2011 έγινε η πρώτη εταιρία στην οποία ήταν δυνατή η ανταλλαγή και αγορά bitcoins. Ένα άλλο από τα δημοφιλέστερα κρυπτονομίσματα, το Litecoin, δημιουργήθηκε το 2011 και είχε την δεύτερη μεγαλύτερη χρηματιστηριακή αξία μετά το Bitcoin μέχρι που το ξεπέρασε τον Οκτώβριο του 2014 το Ripple. Το Litecoin αποτελούσε τροποποιημένη εκδοχή του Bitcoin, με αλλαγές που στόχευαν στην βελτιστοποίηση της ταχύτητας των συναλλαγών, με στόχο την χρήση του σε καθημερινές συναλλαγές. Ένα ακόμα σημαντικό νόμισμα, αποτελεί το Peercoin, που ήταν το πρώτο κρυπτονόμισμα που ενοποίησε τόσο αλγορίθμους Proof-of-Work όσο και αλγορίθμους Proof-of-Stake για την επαλήθευση της συναλλαγής και έτσι αποτελεί το πρώτο υβριδικό κρυπτονόμισμα. Το 2013 γεννήθηκε η ιδέα του Ethereum, ενός απομακρυσμένου και ανοιχτού κώδικα Blockchain, που δίνει την δυνατότητα στους χρήστες του να δημιουργήσουν να δημιουργούν απομακρυσμένες εφαρμογές από τον Vitaly Buterin, ενώ έναν χρόνο μετά ξεκίνησε η διαδικασία σχεδίασής του. Στις 30 Ιουλίου του 2015 το δίκτυο ξεκίνησε να λειτουργεί. Το επίσημο κρυπτονόμισμα του Ethereum λέγεται Ether και αποτελεί ένα από τα δημοφιλέστερα κρυπτονομίσματα μαζί με το Bitcoin.

Η μεγαλύτερη κριτική που δέχθηκαν τα κρυπτονομίσματα είναι η τεράστια δαπάνη ενέργειας, στον αλγόριθμο PoW για την επαλήθευση μίας συναλλαγής. Έτσι μετέπειτα κρυπτονομίσματα προσπάθησαν να λύσουν το πρόβλημα αυτό με την δημιουργία νέων αλγορίθμων επαλήθευσης συναλλαγών που θα οδηγούσαν σε μεγαλύτερη αποδοτικότητα, ταχύτητα και μείωση της καταναλωθείσας ενέργειας. Τον Σεπτέμβριο του 2022 το Ethereum μετέβηκε από την χρήση του PoW στην χρήση του PoS, σαν μηχανισμό ομοφωνίας του και υπολογίστηκε από τον δημιουργό του, ότι η μετάβαση θα μείωνε την παγκόσμια κατανάλωση ηλεκτρικής ενέργειας κατά 0.2%. Οι αποδοτικότεροι αλγόριθμοι αυτή την στιγμή είναι οι αλγόριθμοι τύπου βυζαντινών σφαλμάτων (BFT, PBFT, IBFT, FBA). Αυτοί παρέχουν ταχύτερη επαλήθευση των

συναλλαγών και μικρή κατανάλωση ενέργειας – από την στιγμή που δεν βασίζονται σε δύσκολους μαθηματικούς υπολογισμούς - πράγμα που τους ξεχωρίζει από τους αλγορίθμους PoW και PoS. Παρόλα αυτά αυτοί οι τύποι αλγορίθμων δεν είναι επεκτάσιμοι και είναι δύσκολη η χρήση τους σε δημόσια Blockchain. Τα δύο πιο διάσημα κρυπτονομίσματα που χρησιμοποιούν τον αλγόριθμο FBA είναι τα Ripple και Stellar.

3.2 Τα κυριότερα κρυπτονομίσματα

3.2.1 Bitcoin

Όπως αναφέρθηκε και πριν, το Bitcoin αποτελεί την πρώτη υλοποίηση κρυπτονομισμάτων στην αγορά. Αποτελεί δηλαδή πρόγονο όλων των κρυπτονομισμάτων, και την βάση στην οποία εξελίχθηκε όλη η τεχνολογία που μελετάμε. Ουσιαστικά πρόκειται για ένα ψηφιακό νόμισμα, χωρίς υλική υπόσταση με αποκεντρωμένο χαρακτήρα.

Στόχος του ήταν η χρήση του ως ψηφιακό συνάλλαγμα για ηλεκτρονικές αγορές, χωρίς όμως την εξάρτηση από κάποιο διαμεσολαβητή που θα παρέχει εγγύηση συναλλαγής. Τον στόχο αυτό τον πέτυχε χρησιμοποιώντας την δομή του Blockchain – αποτελώντας την πρώτη πρακτική χρήση της τεχνολογίας αυτής – και με την εισαγωγή της έννοιας του αλγορίθμου ομοφωνίας, ενός αλγορίθμου που απαιτεί τεράστια υπολογιστική ισχύ για να αποτρέπει μεμονωμένους χρήστες από κακόβουλες συμπεριφορές. Ο αλγόριθμος που πρότεινε ήταν το Proof of Work, ο πυλώνας στον οποίο βασίστηκαν και τα υπόλοιπα μεταγενέστερα κρυπτονομίσματα.

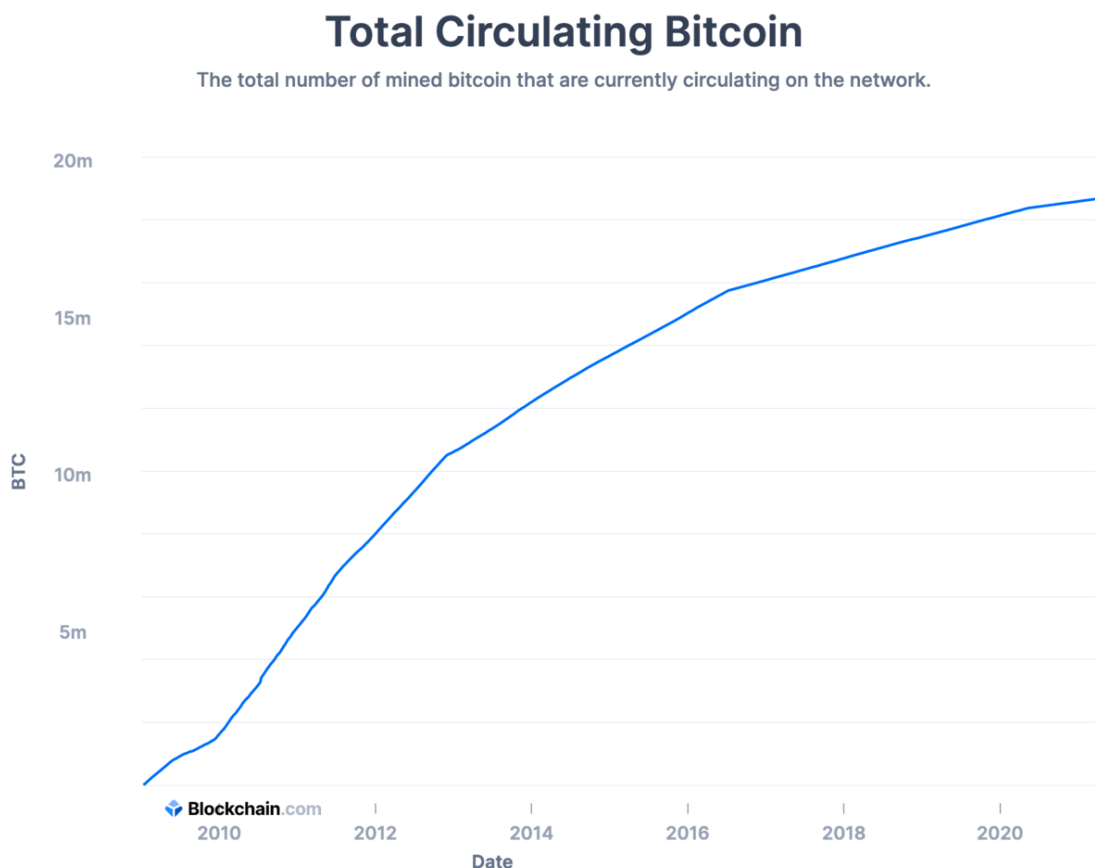
Για να είναι επιτυχής μία τέτοια δομή, θα έπρεπε οι κόμβοι που ξοδεύουν ενέργεια να επιβραβεύονται για την λύση του προβλήματος του Proof of work, με κάποια μορφή συναλλάγματος. Έτσι στην δομή του, το Bitcoin ήταν αυτό που εισήγαγε και την έννοια του miner, ενός κόμβου δηλαδή που όταν προσθέτει στο σύστημα ένα μπλοκ συναλλαγών, επιβραβεύεται με κάποια bitcoins. Η ανταμοιβή για την λύση ενός μπλοκ είναι μεγάλη, και για τον λόγο αυτό οι miners συγκεντρώνονται σε ομάδες για την λύση του προβλήματος από κοινού. Όλες οι παραπάνω προδιαγραφές για το σύστημα δημοσιεύθηκαν από τον Satoshi Nakamoto το 2009 (ενδέχεται να είναι ψευδώνυμο και να αποτελείται από πολλά πρόσωπα). Από τότε η επιστημονική κοινότητα μεγάλωσε εκθετικά το Bitcoin με πολλούς προγραμματιστές, από την στιγμή που το project ήταν ανοιχτού κώδικα.

Έτσι το λογισμικό είναι ανοιχτό προς όποιον προγραμματιστή θέλει να επιθεωρήσει τον κώδικα ή να φτιάξει τις δικές του τροποποιημένες εκδόσεις του Bitcoin. Αυτό ήταν και ο σημαντικότερος παράγοντας για την ραγδαία ανάπτυξη πολλών και διάφορων κρυπτονομισμάτων, καθώς οι δημιουργοί τους εμπνεύστηκαν και δανείστηκαν στοιχεία από το ίδιο το Bitcoin προσπαθώντας να το βελτιώσουν.

Η χρήση του λογισμικού είναι δωρεάν και διαθέσιμη σε όλες τις χώρες του κόσμου, εφόσον υπάρχει σύνδεση στο διαδίκτυο. Η βασική λειτουργία του λογισμικού έγκειται στην εκτέλεση συναλλαγών bitcoins και την αναμετάδοση πληροφοριών ανάμεσα σε κόμβους και την επιβεβαίωση της εγκυρότητάς τους για το υπόλοιπο δίκτυο. Καθώς το λογισμικό είναι ανοιχτού κώδικα, δύνανται να υπάρχουν πάρα πολύ διαφορετικές εκδόσεις και εκδοχές του. Στην ουσία, ο καθένας θα μπορούσε με τις κατάλληλες

ικανότητες να δημιουργήσει ένα αντίστοιχο δίκτυο, αντιγράφοντας σε μεγάλο βαθμό το λογισμικό του Bitcoin, προσθέτοντας ή διαφοροποιώντας με ότι κανόνες επιθυμεί. Κατά αυτήν την έννοια, τα συστατικά στοιχεία του λογισμικού, έχουν δημιουργηθεί συναινετικά από προγραμματιστές, ενσωματώνοντας καινοτομίες διαθέσιμες από άλλα λογισμικά ανοιχτού κώδικα, αλλά και νέα στοιχεία που δεν είχαν εμφανιστεί πριν.

Από την στιγμή που το Bitcoin προσπαθεί να είναι εικονικό συνάλλαγμα, η εισαγωγή νέων bitcoins περιορίζεται και είναι πεπερασμένη σε συνολικό αριθμό. Αυτό επιτυγχάνεται τεχνικά, και είναι μέρος των κανόνων του δικτύου. Ο μέγιστος αριθμός bitcoins είναι προκαθορισμένος και δεν θα ξεπεράσει τα 21.000.000. Σήμερα υπάρχουν περίπου 19.400.000 bitcoins αριθμός ο οποίος αλλάζει κάθε 10 λεπτά με την εισαγωγή νέου μπλοκ στο σύστημα και την επιβράβευση των miners. Για την μείωση του ρυθμού παραγωγής των bitcoins, η επιβράβευση για την λύση ενός μπλοκ υποδιπλασιάζεται κάθε 210.000 μπλοκ, η περίπου κάθε 4 χρόνια (Αnon., 2023). Ο ρυθμός παραγωγής τους ελαττώνεται σταδιακά έως περίπου το 2140 όπου θα παραχθεί το τελευταίο. Αυτή η μέθοδος σε κάποιο βαθμό, προσομοιάζει την πορεία διάθεσης ενός πολύτιμου μετάλλου (άργυρος, χρυσός) στην παγκόσμια αγορά. Αρχικά, η εξόρυξή του είναι εύκολη και σχετικά μεγάλες ποσότητες είναι πιο εύκολα διαθέσιμες, προοδευτικά όμως γίνεται σπανιότερο έως ότου εξαντληθούν τα αποθέματα του πλανήτη. Ο ρυθμός της μέχρι τώρα παραγωγής φαίνεται παρακάτω:



Εικόνα 2: Συνολική παραγωγή Bitcoin. Total Circulation of Bitcoin (2023). Blockchain.com

Τα βασικά πλεονεκτήματα του Bitcoin είναι τα εξής:

- **Ταχύτητα συναλλαγών/ Διεθνής φύση:** Οι συναλλαγές σε bitcoin γίνονται άμεσα και ενημερώνονται ταυτόχρονα σε όλο το δίκτυο παγκοσμίως. Το μόνο που απαιτείται είναι σύνδεση στο διαδίκτυο.
- **Έλεγχος από τον χρήστη:** Ο χρήστης είναι ο μόνος που έχει την δυνατότητα εκτέλεσης συναλλαγής και εφόσον δεν έχει παραχωρήσει αυτό το δικαίωμα, και έχει προστατεύσει λογικά την πρόσβαση στα bitcoins του, είναι πρακτικά αδύνατο να κλαπούν από τρίτους (εφόσον η κρυπτογράφηση δεν παραβιαστεί).
- **Διαφάνεια συναλλαγών:** Όλες οι συναλλαγές καταγράφονται στο σύστημα και είναι διαθέσιμες δημόσια για όλους τους χρήστες. Έτσι κάθε χρήστης μπορεί να εξετάσει οποιαδήποτε διεύθυνση και να δει όλες τις συναλλαγές που έχουν εκτελεστεί με αυτή.
- **Συναινετική φύση αλλαγών:** Η αλλαγή οιαδήποτε χαρακτηριστικού του λογισμικού ή των κανόνων του, έχει ουσιαστικά εφαρμογή μόνο όταν τις δεχτεί η κοινότητα που απαρτίζει το δίκτυο. Με αυτό τον τρόπο αποφεύγονται κακόβουλες αλλαγές που θα μπορούσαν να αλλάξουν θεμελιωδώς το λογισμικό (καθώς η πλειοψηφία των χρηστών θα τις αναγνωρίσει και δεν θα τις δεχτεί), αλλά και μεγάλη ευελιξία και ταχύτητα αντίδρασης σε περίπτωση εντοπισμού σφαλμάτων ή απρόβλεπτων αστοχιών κατά τη λειτουργία.
- **Αποκεντρωμένη φύση:** Το θεμέλιο του Bitcoin ήταν η αποκεντρωμένη φύση και το να μην χρειάζεται κάποια κεντρική αρχή για την επιβεβαίωση των συναλλαγών πράγμα που έχει πετύχει. Η μόνη πιθανότητα αποτυχίας αυτού είναι το 51% attack, όπου ένας συγκεκριμένος χρήστης έχει στην κατοχή του περισσότερο από το 50% των υπάρχοντων miner του συστήματος, όπου και θα μπορούσε να τροποποιήσει το Blockchain όπως αυτός επιθυμεί.
- **Μη αντιστρέψιμη φύση:** Από την στιγμή που επαληθευτεί μία συναλλαγή στο Bitcoin, δεν μπορεί να αναιρεθεί, γιατί η ιστορικότητα του Blockchain δεν μπορεί να αλλάξει. Αυτό έχει το επιπλέον πλεονέκτημα προς όσους διαθέτουν προϊόντα για bitcoin ότι δεν είναι δυνατόν να ανακληθούν συναλλαγές όπως πχ είθισται στις απάτες με πιστωτικές κάρτες.

Παρακάτω φαίνεται η ιστορική τιμή του Bitcoin τα τελευταία χρόνια:



Εικόνα 3: Ιστορική τιμή του Bitcoin από το 2018. Historical Bitcoin price (2023). Blockchain.com

3.2.2 Litecoin

Το Litecoin είναι από τα πρώτα κρυπτονομίσματα που εμπνεύστηκαν το Bitcoin υιοθετώντας την λογική του. Δημιουργήθηκε το 2011 από τον Charlie Lee, έναν πρώην υπάλληλο της Google. Στόχος του ήταν να είναι μία “lite” εκδοχή του Bitcoin χρησιμοποιώντας τα ίδια χαρακτηριστικά αλλά όντας “ελαφρύτερο”. Συνήθως αναφέρεται και σαν ψηφιακό ασήμι, όπου το Bitcoin αναφέρεται σαν ψηφιακός χρυσός. Τεχνικά το Litecoin μοιράζεται και τεχνικά πολλά κοινά με το Bitcoin, καθώς οι δημιουργοί του χρησιμοποίησαν σαν βάση τον κώδικα του Bitcoin. Έτσι πρόκειται, όπως και το Bitcoin, για ένα peer-to-peer κρυπτονομίσμα ανοιχτού κώδικα το οποίο χρησιμοποιεί και αυτό έναν τύπο αλγορίθμου Proof of work για την επαλήθευση των συναλλαγών του. Η διαφορά με το Bitcoin είναι ότι ο αλγόριθμος κρυπτογράφησης που χρησιμοποιεί ονομάζεται Scrypt και είναι γρηγορότερος από τον SHA-256, που χρησιμοποιεί το Bitcoin, προσφέροντας στην πλατφόρμα γρηγορότερη επιβεβαίωση συναλλαγών. Στο Litecoin τα μπλοκ εισέρχονται στο σύστημα κάθε 2.5 λεπτά σε αντίθεση με τα 10 που παίρνει στο Bitcoin.

Λόγω την κοινής σύνδεσης που έχουν τα δύο μεταξύ το Litecoin χρησιμοποιείται σαν ένα περιβάλλον δοκιμής νέων χαρακτηριστικών για την μετέπειτα εισαγωγή τους στο Bitcoin. Παράλληλα αποτελεί το δεύτερο δημοφιλέστερο νόμισμα σε αριθμό miners. Όπως και το Bitcoin το Litecoin έχει έναν περιορισμένο αριθμό coin που θα παραχθούν και χρησιμοποιεί την ίδια υποδιαίρεση των επιβραβεύσεων για την λειτουργία της οικονομίας του.

Οι διαφορές που αυτό έχει με το Bitcoin είναι οι εξής:

- Όπως ήδη αναφέρθηκε το Litecoin επιβεβαιώνει τις συναλλαγές του 4 φορές γρηγορότερα, παρέχοντας μεγαλύτερη απόκριση για τους χρήστες του.
- Ο αλγόριθμος κρυπτογράφησης της μορφής του PoW που χρησιμοποιεί ονομάζεται Scrypt και βασίζεται στην χρήση την μνήμης του υπολογιστή για την δυσκολία εύρεσης της λύσης. Ο αλγόριθμος αυτός κάνει δυσκολότερη και πιο ακριβή την παραγωγή ειδικών συσκευών (FPGAs και ASIC) για την εξόρυξη του νομίσματος (Coventry, 2012).
- Εξορύσσεται ομαδικά μαζί με ένα άλλο διάσημο κρυπτονομίσμα που ονομάζεται Dogecoin, αυξάνοντας την ανταμοιβή των miners και την ασφάλεια των δύο συστημάτων (McGleenon, 2022).
- Ο μέγιστος αριθμός νομισμάτων που μπορούν να υπάρχουν στο σύστημα είναι 84.000.000 που είναι 4 φορές μεγαλύτερος από αυτόν του Bitcoin.
- Όμοια με το Bitcoin και τα δύο κρυπτονομίσματα αυξάνουν την δυσκολία εξόρυξης μπλοκ κάθε 2016 μπλοκ, αλλά επειδή η ταχύτητα εξόρυξης είναι 4 φορές μεγαλύτερη η αλλαγή της δυσκολίας στο Litecoin συμβαίνει κάθε 3.5 μέρες ενώ στο Bitcoin είναι περίπου 14 μέρες.

Παρακάτω φαίνεται η ιστορική τιμή του Litecoin τα τελευταία χρόνια:

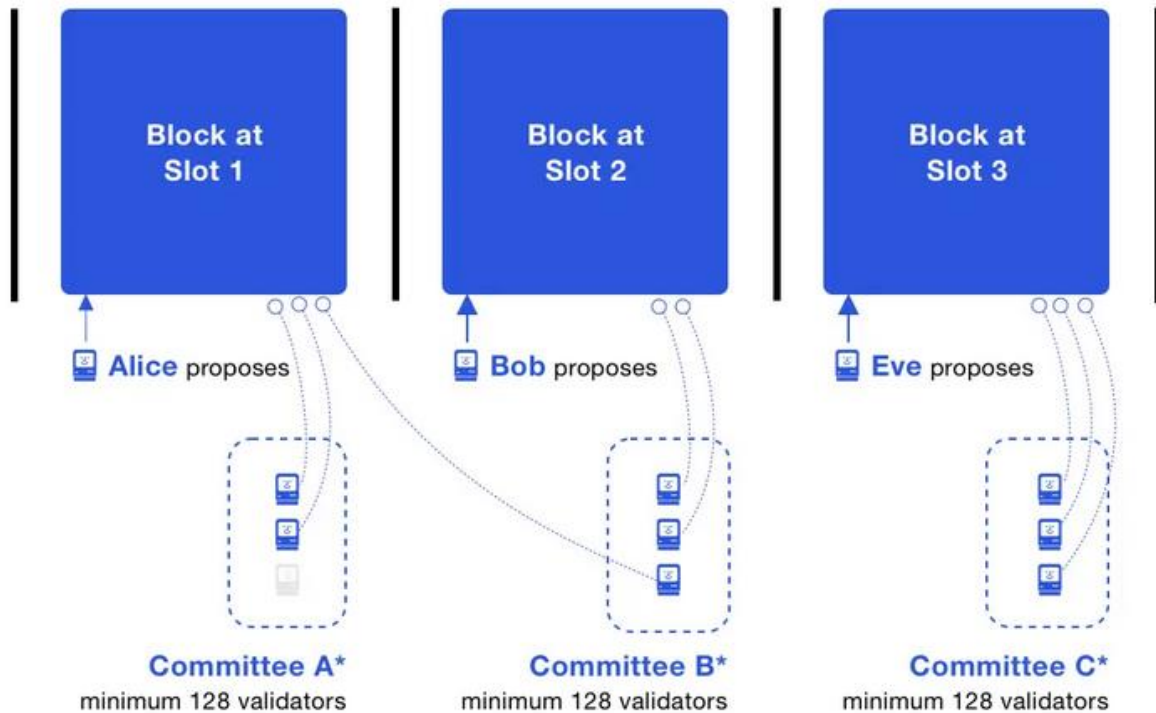


Εικόνα 4: Ιστορική τιμή του Litecoin από το 2018. Historical Litecoin price (2023). Blockchain.com

3.2.3 Ethereum

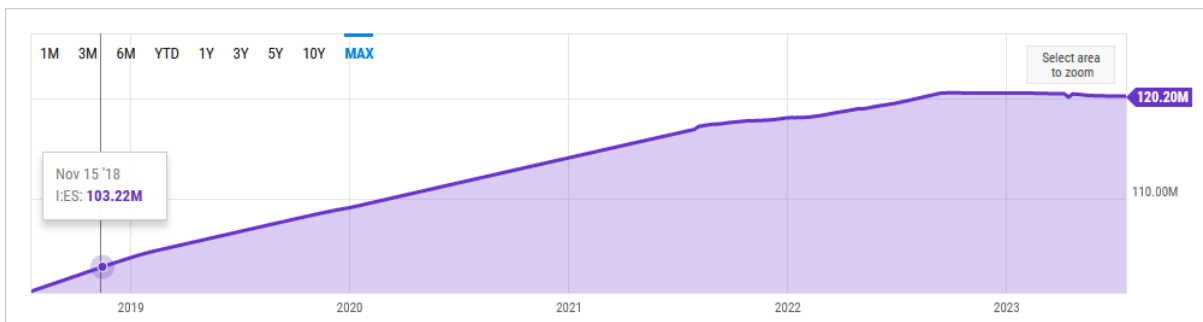
Η αρχική ιδέα του Ethereum εισάχθηκε το 2013 από τον Vitaly Buterin με την δημοσίευση των πρώτων προδιαγραφών του. Το Ethereum, η πλατφόρμα στην οποία λειτουργεί το κρυπτονόμισμα Ether (ETH), προσφέρει στους χρήστες την δυνατότητα δημιουργίας αποκεντρωμένων εφαρμογών οι οποίες βασίζονται σε έτοιμη υλοποίηση του Blockchain από το ίδιο το Ethereum. Η λειτουργία αυτή παρέχει σε πολλούς οργανισμούς και χρήστες την δημιουργία αποκεντρωμένων νομισμάτων, κρυπτονομισμάτων, και non-fungible tokens (NFTs) χωρίς να βασίζονται σε κεντρικές αρχές όπως τράπεζες ή μεσίτες.

Το Ether αποτελεί το κρυπτονόμισμα που παράγεται με βάση τους κανόνες της πλατφόρμας του Ethereum, και το genesis block δημιουργήθηκε το 2015. Το Ether μπορεί να μεταφέρεται μεταξύ λογαριασμών και να χρησιμοποιηθεί για την αντιστάθμιση συμμετεχόντων κόμβων εξόρυξης για τους εκτελούμενους υπολογισμούς, αποτελώντας την αμοιβή των miners. Το Ether αρχικά είχε σαν αλγόριθμο ομοφωνίας έναν τύπο PoW, αλλά το 2018 μετέβηκε σε αλγόριθμο Proof of stake, μειώνοντας τις ενεργειακές του απαιτήσεις κατά 99% (Clark, 2022). Μετά την μεταβίβαση στον αλγόριθμο PoS το Ether χρησιμοποιεί για την επαλήθευσή συναλλαγών της λεγόμενες εποχές (epochs). Το Blockchain χωρίζεται σε θέσεις (slots), όπου η κάθε μία είναι 12 δευτερόλεπτα και κάθε εποχή είναι 32 θέσεις, άρα 6.4 λεπτά. Κάθε θέση αναπαριστά την πιθανότητα ενός μπλοκ να προστεθεί στο Blockchain, και όταν το σύστημα λειτουργεί ομαλά κάθε θέση καλύπτεται και έτσι προστίθεται ένα μπλοκ κάθε 12 δευτερόλεπτα. Παρόλα αυτά υπάρχουν θέσεις που μπορεί να είναι άδειες. Σε κάθε εποχή, οι κόμβοι που πρέπει να επαληθεύσουν τα μπλοκ (validators) επιλέγονται τυχαία και ψηφίζουν για τα μπλοκ που θεωρούν ως σωστά. Κάθε validator έχει τον ρόλο του να επαληθεύσει τα μπλοκ, αλλά παράλληλα να ανιχνεύσει και κακόβουλες συμπεριφορές από άλλους validators. Τελικά ένα μπλοκ θεωρείται επαληθευμένο στο τέλος κάθε εποχής αν έχει επαληθευτεί από τα 2/3 των ενεργών validators (ethos.dev, 2022).



Εικόνα 5: Ψηφοφορία των validators για το ποιο μπλοκ θεωρούν σωστό. Validator committee (2023). ethos.dev

Στο τέλος κάθε εποχής νέα ETH δημιουργούνται και δίνονται στους validators. Η παραγωγή του ETH είναι θεωρητικά άπειρη, παρόλα αυτά η καμπύλη πληθωρισμού της φαίνεται να μειώνεται εκθετικά σε σχέση με τον χρόνο. Έκτος από την επιβράβευση των miners με την δημιουργία καινούργιου ETH, οι χρήστες πληρώνουν τους validators μία προμήθεια που ονομάζεται “gas”, για να προωθείται η εξόρυξη και η ασφάλεια του συστήματος.



Εικόνα 6: Συνολική παραγωγή ETH. Total Ethereum Circulation (2023). Blockchain.com

Κάποιες διαφορές του Ethereum με το Bitcoin είναι:

- Ο μέγιστος αριθμός νομισμάτων που μπορούν να υπάρχουν στο Ethereum είναι άπειρος σε αντίθεση με τις πεπερασμένες τιμές στα Bitcoin και Litecoin.
- Η προμήθεια για μία συναλλαγή στο Ethereum πληρώνεται από τον κόμβο που θέλει να κάνει την συναλλαγή σε αντίθεση με το Bitcoin, όπου η προμήθεια πληρώνεται από το ίδιο το σύστημα.
- Το Ethereum χρησιμοποιεί τον αλγόριθμο PoS που μειώνει σημαντικά τις απαιτήσεις σε ενέργεια σε σχέση με το Bitcoin.

- Το Ethereum έχει εισάγει την τεχνολογία των έξυπνων συμβολαίων σε αντίθεση με το Bitcoin.

Στις αρχές του 2018 το Ether έγινε το δεύτερο ακριβότερο κρυπτονόμισμα με βάση την χρηματιστηριακή του αξία. Το Ether έφτασε την υψηλότερη τιμή του τον Νοέμβριο του 2021 με αξία περίπου ίση με 4800 δολάρια ανά μονάδα. Η πορεία της τιμής του είναι παρόμοια με του Bitcoin, πράγμα που φαίνεται και από την γραφική αναπαράσταση της πορείας της:



Εικόνα 7: Ιστορική τιμή του Ethereum από το 2018. Historical Ethereum price (2023). Blockchain.com

3.3 Εμπορία & ανταλλαγή κρυπτονομισμάτων (Cryptocurrency trading)

Η εμπορία κρυπτονομισμάτων πραγματοποιείται από συγκεκριμένες επιχειρήσεις που ονομάζονται ανταλλακτήρια κρυπτονομισμάτων. Οι επιχειρήσεις μπορούν να είναι απλά μεσάζοντες που λαμβάνουν μία προμήθεια για την συναλλαγή ή να λειτουργούν θέτοντας μία δημόσια τιμή αγοράς/πώλησης. Ουσιαστικά η δεύτερη μορφή εταιρίας αγοράζει τις μετοχές και τις πουλάει μετά στους πελάτες της. Οι συναλλαγές μπορούν να γίνουν με διαφορετικούς τρόπους. Ένας από αυτούς είναι η μεταφορά των κρυπτονομισμάτων στο πορτοφόλι του χρήστη.

Μετά την δημιουργία του bitcoin το 2008, δημιουργήθηκαν πολλές πλατφόρμες ανταλλαγών ψηφιακών αγαθών, με κύριο στόχο την ανταλλαγή κρυπτονομισμάτων. Η πρώτη εταιρία ήταν η Mt.Gox που ιδρύθηκε το 2010. Τον Φεβρουάριο του 2014 η εταιρία έκλεισε τον ιστότοπο της και σταμάτησε την εμπορική της δραστηριότητα, εξαιτίας υποκλοπής κρυπτονομισμάτων από το ψηφιακό της πορτοφόλι.

Σήμερα υπάρχει αφθονία ανταλλακτηρίων που δραστηριοποιούνται ψηφιακά. Κάποια από τα δημοφιλέστερα αποτελούν τα Binance, Coinbase Exchange, Kraken, KuCoin, Gemini, OKX. Ο όγκος των συναλλαγών από τα ανταλλακτήρια αγγίζει την κλίμακα των δισεκατομμυρίων δολαρίων. Ο δημοφιλέστερος ιστότοπος, είναι το Binance που τον Νοέμβριο του 2022 αγγίζει τα 19 δισεκατομμύρια δολάρια σε συναλλαγές, έχει 15 εκατομμύρια επισκέψεις χρηστών ανά βδομάδα και ανταλλάσσει 386 διαφορετικά κρυπτονομίσματα. Δίνονται και τα δημοφιλέστερα ανταλλακτήρια με βάση το coinmarketcap.com (Νοέμβριος 2022).

Όνομα	Μέγεθος αγοράς (24ω)	Εβδομαδιαίες επισκέψεις	Πλήθος νομισμάτων
Binance	\$19,116,295,700	15,046,498	386
Coinbase Exchange	\$2,228,866,146	959,236	229
Kraken	\$784,253,046	990,352	217
KuCoin	\$758,420,926	1,916,603	763
Binance.US	\$326,498,230	433,833	145
Bitfinex	\$250,334,896	446,991	174
Gemini	\$69,758,824	247,069	112
Coincheck	\$77,053,447	281,696	2
Bitstamp	\$319,307,532	239,341	73
Bybit	\$791,083,721	3,620,694	329
OKX	\$1,315,976,903	1,460,509	352
BitFlyer	\$94,909,174	402,066	9
Bitget	\$801,686,822	1,430,973	400
Bithump	\$420,332,813	328,288	193
Gate.io	\$1,142,238,680	1,588,202	1553

Πίνακας 1: Τα δημοφιλέστερα ανταλλακτήρια κρυπτονομισμάτων. Νοέμβριος 2022. Coinmarketcap.com

Παρακάτω φαίνεται η χρηματιστηριακή αξία, καθώς και άλλα στοιχειώδη χαρακτηριστικά κάποιων δημοφιλών κρυπτονομισμάτων με βάση το coinmarketcap.com. Είναι εύκολα κατανοητό ότι οι χρήστες που ανταλλάσσουν κρυπτονομίσματα προτιμούν να μην τα αγοράζουν και τα φυλάσσουν στα δικά τους πορτοφόλια, αλλά να χρησιμοποιούν τους διαμεσολαβητές που είναι τα ανταλλακτήρια.

Δημιουργία	Νόμισμα	Χρημ. Αξία (Νοε 2022)	Αλγόριθμος Hash	Μηχανισμός Ομοφωνίας
Ιαν 2009	Bitcoin	\$317,720,575,245	SHA-256	PoW
Ιουλ 2015	Ether	\$150,945,699,282	SHA-256	PoS
Δεκ 2013	Dogecoin	\$11,316,865,737	Scrypt	PoW
Μαι 2020	Polkadot	\$6,536,127,992	Blake2b	NPoS
Οκτ 2011	Litecoin	\$4,048,122,532	Scrypt	PoW
Ιουλ 2014	Stellar	\$2,245,919,975	ECDSA	Byzantine Consensus

Πίνακας 2: Κάποια δημοφιλή κρυπτονομίσματα. Νοέμβριος 2022. Coinmarketcap.com

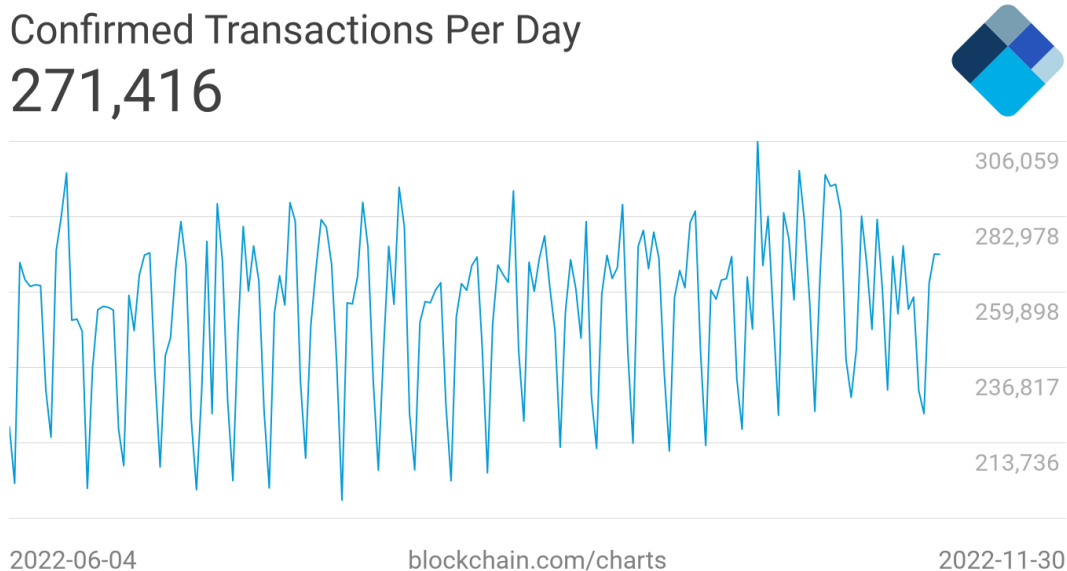
3.4 Επένδυση σε κρυπτονομίσματα

Όπως και τα συμβατικά νομίσματα, έτσι και τα κρυπτονομίσματα, έχουν την αξία που τους δίνουν οι χρήστες. Ένα κρυπτονόμισμα με λίγους χρήστες, πρακτικά δεν είναι νόμισμα γιατί δεν είναι αποδεκτό σαν αντάλλαγμα στο μεγαλύτερο πλήθος συναλλαγών. Έτσι δημιουργείται το πρόβλημα του ότι ένα κρυπτονόμισμα που δεν είναι καθολικό, πρακτικά δεν έχει κάποια αξία διαφορετική από αυτή που ορίζουν οι ίδιοι οι χρήστες. Σε περίπτωση που οι χρήστες σταματήσουν να ενδιαφέρονται για το συγκεκριμένο νόμισμα, η αξία του ξαφνικά θα χαθεί, πράγμα που κάνει αρκετά ασταθή την τιμή του οποιουδήποτε κρυπτονομίσματος. Στα πλαίσια του σήμερα, ακόμα και το δημοφιλέστερο κρυπτονόμισμα – το Bitcoin – δεν είναι αρκετά δημοφιλές και είναι αποδεκτό σε πολύ μικρό αριθμό εμπόρων. Έτσι ο κύριος παράγοντας που επηρεάζει την αξία ενός κρυπτονομίσματος είναι το πλήθος των χρηστών του. Σε αυτή την αστάθεια συμβάλει επίσης το γεγονός ότι η χρήση των κρυπτονομισμάτων σε συναλλαγές είναι πολύπλοκη και αργή. Ήδη αναφέρθηκε ότι οι χρήστες προτιμούν να μην έχουν στα προσωπικά τους πορτοφόλια κρυπτονομίσματα, καθώς η διαχείρισή τους είναι αρκετά πιο δύσκολη σε σχέση με ένα συμβατικό πορτοφόλι. Αυτό έχει σαν αποτέλεσμα οι κύριες συναλλαγές που γίνονται να αφορούν τις συναλλαγές που πραγματοποιούν τα διάφορα ανταλλακτήρια σε μορφή επένδυσης. Ένας ακόμα λόγος που προσδίδει αστάθεια είναι η ευκολία «διαγραφής» του ψηφιακού πλούτου σε κρυπτονομίσματα. Μία αναπάντεχη αστοχία στο σύστημα ή μία πιθανή επίθεση από hackers μπορεί πρακτικά να εξαλείψει τα στοιχεία οποιουδήποτε ψηφιακού πορτοφολιού. Σε αυτό το κομμάτι, σε σχέση με τα τραπεζικά συστήματα, η ασφάλεια και οι άνθρωποι πίσω από ένα κρυπτονόμισμα είναι σημαντικά λιγότεροι και πρακτικά η πιθανότητα κάποιου «περιστατικού» αυξάνεται. Όλα τα παραπάνω, προσδίδουν ένα μεγάλο ρίσκο όσον αφορά τις επενδύσεις σε κρυπτονομίσματα καθώς η αγορά που τα αφορά είναι εξαιρετικά ριψοκίνδυνη και επηρεάζεται από εκατοντάδες παράγοντες.

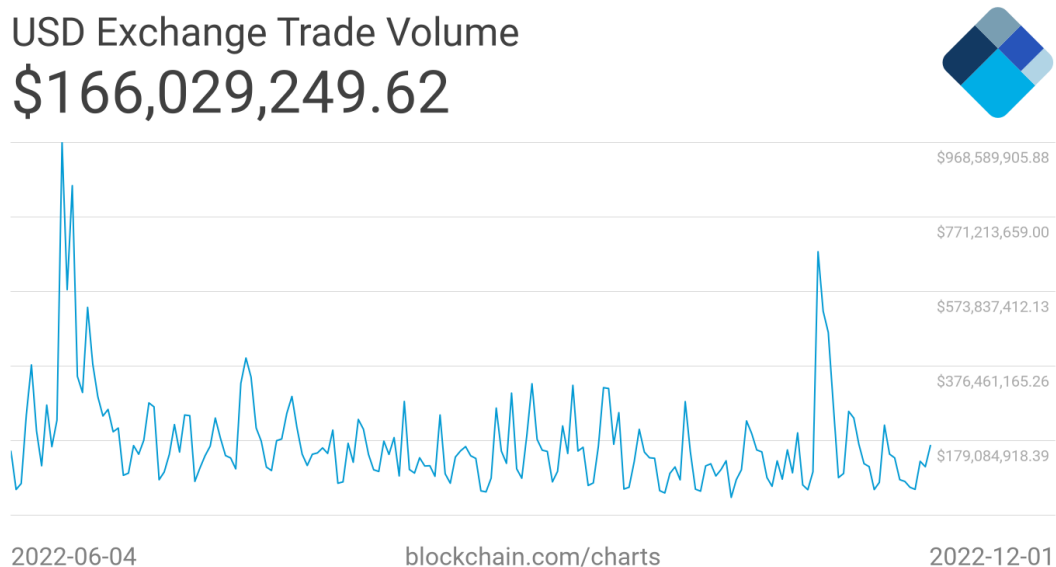
Γενικά, η αστάθεια της τιμής των κρυπτονομισμάτων μπορεί να φανεί εύκολα βλέποντας δεδομένα παλαιών τιμών τους. Οι γραφικές παραστάσεις που παρατέθηκαν πριν δείχνουν μεγάλες πτώσεις της τιμής των δύο δημοφιλέστερων κρυπτονομισμάτων – των Bitcoin και Ether - περίπου τις ίδιες περιόδους, και αναφορικά στις αρχές του 2018, στα μέσα και τέλη του 2021, όπως και στα μέσα του 2022. Η τωρινή περίοδος αποτελεί την χειρότερη πτώση που έχει καταγραφεί προς το παρόν για τα κρυπτονομίσματα και οφείλεται σε πολλούς παράγοντες, όπου ένας από τους σημαντικότερους είναι η ενεργειακή κρίση που κάνει την εξόρυξη κρυπτονομισμάτων μη κερδοφόρα, λόγω της ραγδαίας αύξησης της ηλεκτρικής ενέργειας. Παρόλα αυτά ο αριθμός των συναλλαγών και ο όγκος των συναλλαγών δεν έχει δει μείωση το τελευταίο εξάμηνο, πράγμα που σημαίνει ότι το επενδυτικό ενδιαφέρον παραμένει μεγάλο.

Ο κυριότερος παράγοντας που κάνει τα κρυπτονομίσματα ποθητή επένδυση για πολλούς είναι το περιθώριο κέρδους που μπορεί να προκύψει σε μικρό χρονικό διάστημα. Οι κλασικές επενδυτικές επιλογές που θεωρούνται ασφαλείς μπορεί να προσφέρουν αυξήσεις της τάξης του 5-10% τον χρόνο, και το κέρδος από μία τέτοια κίνηση φαίνεται σε βάθος χρόνου. Από την άλλη ένα επιτυχημένο κρυπτονόμισμα όπως το Bitcoin, σε ανθίζουσες περιόδους, μπορεί να έβλεπε αύξηση της τιμής του πάνω από το 100% της αξίας του. Γι' αυτό κυρίως τον λόγο, παρόλο που οι συμμετέχοντες

γνωρίζουν το πλήθος των περιορισμών της αγοράς των κρυπτονομισμάτων, φαίνεται να μην δίνουν σημασία στην επικινδυνότητα αυτών των συναλλαγών. Επίσης μεγάλο ποσοστό των ανθρώπων που επενδύουν φαίνεται να προέρχεται από κερδοσκοπική φρενίτιδα, λόγω της δημοσιότητας και της προώθησης των κρυπτονομισμάτων από άλλα μέσα. Έτσι, είναι εμφανές ότι τα κρυπτονομίσματα αποτελούν ακόμα και τώρα ποθητή επένδυση από πολλούς, πράγμα που φαίνεται και από τον συνολικό αριθμό συναλλαγών τους, ο οποίος παραμένει σχετικά σταθερός.



Εικόνα 8: Επαληθευμένες συναλλαγές του Bitcoin ανά μέρα. Confirmed Transactions Per Day (2022). Blockchain.com



Εικόνα 9: Όγκος συναλλαγών Bitcoin σε USD. USD Exchange Trade Volume (2022). Blockchain.com

3.5 Σύντομη βιβλιογραφική επισκόπηση

Είναι γενικά αποδεκτό ότι τα κρυπτονομίσματα είναι αρκετά διαφορετικά από τις κλασικές μορφές συναλλάγματος. Κάποιες έρευνες ασχολούνται με την σχέση μεταξύ

των κρυπτονομισμάτων και μεταξύ άλλων επενδυτικών αγαθών, κυρίως με κύριο πυλώνα σύγκρισης το Bitcoin (Baur, et al., 2018; Corbet, et al., 2019; Corbet, et al., 2018). Ο Baur καταλήγει στο ότι το Bitcoin διαφέρει σε κέρδη με τα επενδυτικά αγαθά που συγκρίνεται, και ότι μπορεί να παρέχει αρκετά πλεονεκτήματα για επενδύσεις σε σχέση με τα συμβατικά νομίσματα τόσο σε περιόδους ευστάθειας όσο και αστάθειας της αγοράς. Επίσης τονίζει ότι περίπου το ένα τρίτο των Bitcoins που υπάρχουν στον ledger κατέχονται από επενδυτές που δεν χρησιμοποιούν ποτέ το νόμισμα ως ανταλλακτήριο αλλά κρατάνε τα ποσά τους για επενδυτικούς σκοπούς. Παρόμοια ο Corbet καταλήγει στο ότι τα κρυπτονομίσματα έχουν μεγάλη συσχέτιση μεταξύ τους αλλά μικρή με άλλα συγκρινόμενα επενδυτικά αγαθά και αυτό το γεγονός κάνει τα κρυπτονομίσματα μία καλή επένδυση για την διεύρυνση του επενδυτικού portfolio, σε σχέση με τα συμβατικά νομίσματα. Παρόλα αυτά τονίζεται ότι εξαιτίας της αστάθειας των κρυπτονομισμάτων και την έλλειψη ερευνών για τους παράγοντες που τα επηρεάζουν, αυξάνεται η πιθανότητα κακών επενδύσεων.

Σε συμφωνία με τα προηγούμενα, τα κρυπτονομίσματα φαίνεται να έχουν πολύ μικρή, έως καθόλου συσχέτιση με πολλά βασικά αγαθά: πετρέλαιο, λευκόχρυσος, χρυσός και σιτάρι (Lawuobahsumo, et al., 2022). Παρόλα αυτά, ειδικά για τον χρυσό, ο Lawuobahsumo τονίζει ότι παρουσιάζεται αρνητική συσχέτιση σε σχέση με τα κρυπτονομίσματα. Συνεπώς οι παράγοντες που επηρεάζουν την τιμή των κρυπτονομισμάτων φαίνεται, τις περισσότερες φορές, να είναι διαφορετικοί από τους παράγοντες που επηρεάζουν τα παραπάνω βασικά αγαθά. Από την άλλη μελετήθηκε η συσχέτιση της απότομης πτώσης της τιμής του πετρέλαιο με τα κρυπτονομίσματα (Kumah & Jones, 2022), αλλά τα αποτελέσματα που αντλήθηκαν δεν καταλήγουν σε ένα στιβαρό συμπέρασμα, καθώς η συσχέτιση που βρέθηκε μεταξύ του πετρελαίου και των κρυπτονομισμάτων αφορά μόνο τα 4 από τα 7 κρυπτονομίσματα που μελετήθηκαν και μελετά μόνο ακραίες πτώσεις των τιμών τους, που μπορεί να δικαιολογηθεί από μικρό κραχ της αγοράς που επηρεάζει όλες τις επενδύσεις.

Παράλληλα, έχει μελετηθεί η σχέση τους με την ενέργεια που αυτά καταναλώνουν (Maiti, 2022), χωρίς όμως να αντλείται κάποιο τελικό συμπέρασμα για την σχέση τους. Η έρευνα βρίσκει υψηλή συσχέτιση με την καταναλωθείσα ενέργεια, αλλά μόνο για συγκεκριμένα εύρη τιμών, με χαμηλή συσχέτιση σε άλλα εύρη. Έχει επίσης μελετηθεί και η σχέση της καταναλωθείσας ενέργειας με τον όγκο συναλλαγής των κρυπτονομισμάτων (Söylemez & Gürsoy, 2022), χωρίς να βρεθεί κάποια συσχέτιση. Επιπλέον η έρευνα βρίσκει ότι η πτώση της καταναλωθείσας ενέργειας προκαλεί και πτώση τιμής του Bitcoin, χωρίς όμως να ισχύει το αντίθετο σε περίπτωση αύξησης της ενέργειας. Συνεπώς και αυτή η έρευνα παρατηρεί συσχέτιση της ενέργειας και της τιμής των κρυπτονομισμάτων, αλλά μόνο κάτω από συγκεκριμένες προϋποθέσεις. Ένα ακόμα μέγεθος που έχει μελετηθεί ως παράγοντας επιρροής της τιμής των κρυπτονομισμάτων είναι ο όγκος συναλλαγών (Balcilar, et al., 2017) όπου βρέθηκε πως η τιμή ακολουθούν μη γραμμική σχέση με γενικό συμπέρασμα ότι ο όγκος συναλλαγών δεν επηρεάζει πλήρως και με τρόπο ανάλογο την τιμή. Ο όγκος συναλλαγών μπορεί να αποτελέσει σημείο συσχέτισης για την τιμή, αλλά μόνο σε περιπτώσεις που η αγορά βρίσκεται σε «φυσιολογική» μορφή, σε ένα περιβάλλον ευστάθειας, χωρίς ακραίες τιμές, όπως αναφέρει η έρευνα. Αυτό πρακτικά σημαίνει ότι έντονες αυξομειώσεις στην αγορά καθώς και κατά την διάρκεια που η αγορά βρίσκεται στα χαμηλά ή υψηλά

σημεία της, ο όγκος συναλλαγών θα παράξει ανεπαρκή συμπεράσματα για την πρόβλεψη της τιμής των κρυπτονομισμάτων.

Επίσης σε άλλες έρευνες μελετήθηκε η σχέση του Bitcoin με συμβατικά νομίσματα όπου προκύπτει πως η αστάθεια των κρυπτονομισμάτων είναι 6-7 φορές μεγαλύτερη από οποιοδήποτε G10 νόμισμα (Osterrieder & Lorenz, 2017). Παρόμοια υπάρχουν και έρευνες για την συσχέτιση μεταξύ κρυπτονομισμάτων (Chu, et al., 2017; Ciaian, et al., 2018), οι οποίες διαπιστώνουν ότι υπάρχει μικρό επίπεδο συσχέτισης στα κύρια κρυπτονομίσματα που μελετήθηκαν. Τα αποτελέσματα της πρώτης έρευνας είναι περιορισμένου χρόνου (Ιούνιος 2014 – Σεπτέμβριος 2016) και το πλήθος των κρυπτονομισμάτων που μελετώνται είναι μικρό, ενώ στην δεύτερη το πλήθος των κρυπτονομισμάτων είναι σαφώς μεγαλύτερο και καταλήγει στο ότι η αγορά του Bitcoin και των Altcoins είναι στενά συνδεδεμένες μεταξύ τους, τόσο βραχυπρόθεσμα όσο και μακροπρόθεσμα. Σε μακροσκοπικό επίπεδο και βραχυπρόθεσμα παρατηρήθηκε μεγάλη συσχέτιση καθώς πτώσεις στην τιμή του Bitcoin επηρέασαν τα 15 από τα 16 altcoins που μελετήθηκαν στην έρευνα. Από την άλλη, για μεγάλο χρονικό διάστημα παρατηρήθηκε ότι ο αριθμός των altcoins που επηρεάζονται είναι σημαντικά χαμηλότερος, αλλά συνεχίζει να υπάρχει σημαντική συσχέτιση σε αυτά που έχουν τα περισσότερα κοινά σε λειτουργία και μορφή με το Bitcoin. Παρόλα αυτά, στην έρευνα τονίζεται ότι η επιτυχία του Bitcoin δεν μπορεί να βοηθήσει στην πρόβλεψη των τιμών των altcoins σε βάθος χρόνου, καθώς είναι απίθανο η υψηλή τιμή του Bitcoin να οδηγήσει σε μακροχρόνια αύξηση της τιμής όλων των κρυπτονομισμάτων πράγμα που μπορεί να δικαιολογήσει την έλλειψη μακροχρόνιας συσχέτισης. Σε συνέχεια της έρευνας του Osterrieder, μελετήθηκε για μεγαλύτερο χρονικό διάστημα η σχέση μεταξύ των 10 μεγαλύτερων κρυπτονομισμάτων (Gkillas, et al., 2018). Η έρευνα βρίσκει μεγάλη συσχέτιση των κρυπτονομισμάτων μεταξύ τους και προσδίδει την συσχέτιση αυτή στις τάσεις που υπάρχουν στην γενική αγορά των κρυπτονομισμάτων.

Συνοψίζοντας, υπάρχουν σίγουρα κάποιες μελέτες για την πρόβλεψη τιμής των κρυπτονομισμάτων και την σύνδεση αυτών με άλλους παράγοντες, αλλά η συσχέτιση των κρυπτονομισμάτων κατά μία μεγάλη περίοδο αστάθειας όπως η σημερινή, που έχει προκαλέσει μείωση τόσο στην δημοτικότητα όσο και στην τιμή των κρυπτονομισμάτων δεν είναι μελετημένη σε βάθος.

Σε συνέχεια με τα παραπάνω, θα μελετήσουμε τα μεγαλύτερα και πιο διαδεδομένα κρυπτονομίσματα καθώς και κάποια που χρησιμοποιούν διαφορετικούς αλγορίθμους ομοφωνίας καθώς και την συσχέτιση με την καταναλωθείσα ενέργεια.

Κεφάλαιο 4. Μέθοδοι προβλέψεων

Η πρόβλεψη, είναι η διαδικασία ανάλυσης παρελθοντικών δεδομένων για την άντληση συμπερασμάτων για το μέλλον. Ο κύριος πυλώνας των προβλέψεων είναι η αβεβαιότητα και το ρίσκο, και μία καλή πρόβλεψη προσπαθεί να ελαχιστοποιήσει και τους δύο αυτούς παράγοντες. Το σημαντικότερο κομμάτι μίας καλής πρόβλεψης είναι η άντληση σωστών δεδομένων, που ελαχιστοποιούν τους παράγοντες της αβεβαιότητας και της τύχης.

Οι προβλέψεις έχουν ένα τεράστιο εύρος χρήσης σε διαφορετικούς τομείς, τόσο στον επιστημονικό όσο και στο επιχειρησιακό τομέα. Στα πλαίσια μίας επιχείρησης, η πρόβλεψη είναι αναπόσπαστο κομμάτι της διαδικασίας λήψης αποφάσεων για το μέλλον της εταιρίας. Τέτοιες αποφάσεις μπορεί να αποτελούν κάποια επέκταση της επιχείρησης, κάποια επένδυση ή και προβλέψεις που αφορούν την επιτυχία ενός προϊόντος στην αγορά. Η ανάπτυξη της επιχείρησης, λοιπόν, συνδέεται άρρηκτα με την ορθή λήψη αποφάσεων μέσα από προβλέψεις των εκατοντάδων παραγόντων που επηρεάζουν την εταιρία. Από την άλλη προβλέψεις χρησιμοποιούνται και στον χρηματοοικονομικό τομέα για την πρόβλεψη τιμών μετοχών στο χρηματιστήριο.

Χρηματιστηριακές προβλέψεις ονομάζονται οι προβλέψεις που προσπαθούν να αποφανθούν για το μέλλον μίας εταιρίας ή μετοχής, με σκοπό το να βρουν πως οι τωρινές τάσεις και μετρικές θα επηρεάσουν την μελλοντική θέση της εταιρίας. Η δυνατότητα ορθής πρόβλεψης μπορεί να φανεί πολύ επικερδής με την αγορά μετοχών σε χαμηλά θέσεις με την προοπτική ανάπτυξης της τιμής της.

Τέλος οι προβλέψεις χρησιμοποιούνται και στον επιστημονικό τομέα, για την βελτίωση της κοινωνίας σε πολλούς τομείς. Κάποιες σημαντικές επιστημονικές εφαρμογές που χρησιμοποιούν τις προβλέψεις είναι η πρόβλεψη σεισμών, η πρόβλεψη ενεργειακών αναγκών με σκοπό την ενσωμάτωση ανανεώσιμων πηγών ενέργεια για την κάλυψη των αναγκών σε ενέργεια, η πρόβλεψη καιρού στον μετεωρολογικό τομέα και η πρόβλεψη των γενικών τεχνολογικών αναγκών της κοινωνίας. Όλα τα παραπάνω στοιχειοθετούν την σημασία των προβλέψεων σε όλα τα πεδία που αντιμετωπίζουν το ζήτημα της λήψης απόφασης.

4.1 Είδη προβλέψεων

Οι προβλέψεις μπορούν να χωριστούν σε δύο μεγάλες κατηγορίες μεθόδων πρόβλεψης, τις ποιοτικές (qualitative) μεθόδους και τις ποσοτικές (quantitative) μεθόδους.

4.1.1 Ποιοτικές μέθοδοι

Οι ποιοτικές μέθοδοι δεν βασίζονται σε μαθηματικές σχέσεις, αλλά στην υποκειμενική κρίση και χρησιμοποιούνται κυρίως σε περιπτώσεις όπου τα παρελθοντικά δεδομένα δεν είναι επαρκή ή όταν οι τάσεις αλλάζουν συνεχώς. Το μεγαλύτερο πλεονέκτημα τους είναι η δυνατότητα προσθήκης νέων πληροφοριών και η εύκολη αντίληψη της αγοράς. Παρόλα αυτά αποτελούν υποκειμενικά μοντέλα τα οποία βασίζονται σε ανθρώπινες απόψεις για την παραγωγή της τελικής πρόβλεψης. Κάποια είδη ποιοτικών προβλέψεων είναι τα εξής:

- Έρευνα αγοράς

- Γνωμάτευση στελεχών
- Μέθοδος των Δελφών
- Αναλογίες

4.1.2 Ποσοτικές Μέθοδοι

Σε αντίθεση με τις ποιοτικές μεθόδους, οι ποσοτικές μέθοδοι προβλέψεων είναι μαθηματικές μέθοδοι που βασίζονται στην δημιουργία μοντέλων για την επεξεργασία συνήθως μεγαλύτερου όγκου δεδομένων σε σχέση με τις ποιοτικές μεθόδους. Οι μέθοδοι αυτοί είναι αντικειμενικές καθώς βασίζονται στην μαθηματική ανάλυση των δεδομένων και έτσι είναι απαλλαγμένες από ανθρώπινη υποκειμενικότητα και συναίσθημα και εξαρτώνται μόνο από τις υπάρχουσες πληροφορίες. Για να λειτουργήσουν αυτά τα μοντέλα, θα πρέπει τα διαθέσιμα δεδομένα να ακολουθούν το ίδιο πρότυπο συμπεριφοράς στο μέλλον καθώς και να είναι αρκετά για να μην καταλήξουμε σε λάθος συμπεράσματα. Αυτό είναι και το μεγαλύτερο μειονέκτημα τους, όπου σε τέτοιες περιπτώσεις προτιμώνται οι ποιοτικές μέθοδοι προβλέψεων.

Οι ποσοτικές μέθοδοι τελικά χωρίζονται σε δύο υποκατηγορίες, τα μοντέλα χρονοσειρών (time-series models) και τα αιτιοκρατικά μοντέλα (causal models).

4.2 Χρονοσειρές και ποιοτικά τους χαρακτηριστικά

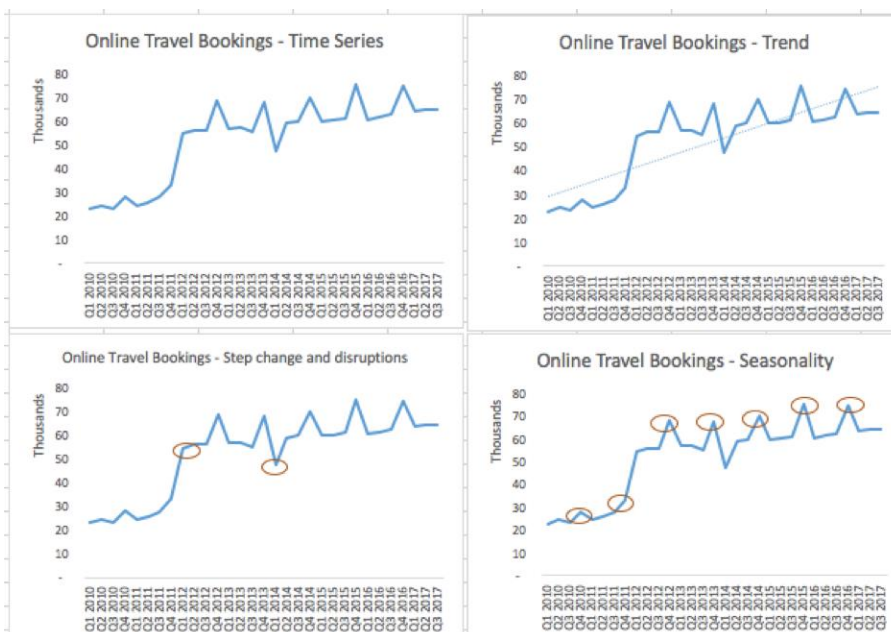
Χρονοσειρά ονομάζεται μία σειρά διακριτών τιμών ενός μεγέθους, που μετρούνται σε ίσα χρονικά διαστήματα. Οι διαδοχικές τιμές της χρονοσειράς δεν είναι απαραίτητα ανεξάρτητες μεταξύ τους και εξαιτίας της εξάρτησης αυτής τα μοντέλα πρόβλεψης χρονοσειρών προσπαθούν να προβλέψουν τις μελλοντικές τιμές. Στην πραγματικότητα, οι μελλοντικές παρατηρήσεις επηρεάζονται από διάφορους παράγοντες όπως τον παράγοντα της τύχης και έτσι δυσκολεύεται η διαδικασία πρόβλεψης. Η σημασία των χρονοσειρών είναι τεράστια καθώς χρησιμοποιούνται σε επιστημονικούς τομείς όπως στην στατιστική, στην αναγνώριση προτύπων, στην επεξεργασία σημάτων, στην πρόβλεψη καιρού, στην οικονομετρία και γενικά σε κάθε τομέα εφαρμοσμένων επιστημών που αφορά χρονικές μετρήσεις και δεδομένα.

Η ανάλυση χρονοσειρών είναι η διαδικασία μελέτης των χρονοσειρών με σκοπό την άντληση χρήσιμων συμπερασμάτων και χαρακτηριστικών των δεδομένων. Οι παράγοντες που μελετάει η ανάλυση χρονοσειρών είναι οι εξής:

- Η **Τάση (Trend)** είναι η μακροπρόθεσμη κατεύθυνση της αλλαγής του μέσου επιπέδου που ακολουθούν τα δεδομένα μίας χρονοσειράς με την πάροδο του χρόνου. Η τάση δείχνει την αύξηση ή την μείωση μίας συμπεριφοράς. Συνήθως οι τάσεις δεν είναι γραμμικές, αλλά με την ανάλυση της χρονοσειράς μπορούν να προκύψουν συμπεράσματα για την γενική πτωτική ή ανοδική πορεία που αυτή ακολουθεί.
- Η **Εποχικότητα (Seasonality)** είναι το φαινόμενο όπου η χρονοσειρά παρουσιάζει την ίδια αναμενόμενη συμπεριφορά που επαναλαμβάνεται κάθε ωρολογιακό έτος. Οποιοδήποτε μοτίβο που επαναλαμβάνεται μέσα σε ένα καθορισμένο χρονικό διάστημα μπορεί να ονομαστεί εποχικότητα. Η εποχικότητα μπορεί να παρατηρείται για διαστήματα μηνών, αλλά συνήθως παρατηρείται μόνο για διαστήματα βδομάδων ή ημερών. Ένα καλό παράδειγμα

εποχιακής συμπεριφοράς είναι τα Χριστούγεννα, όπου αυξάνεται το πλήθος των πωλήσεων εξαιτίας των γιορτών. Η εποχιακή συμπεριφορά εξηγεί ασυνήθιστες αυξήσεις και μειώσεις στα δεδομένα που είναι συστηματικές σε συγκεκριμένα διαστήματα ανά τον χρόνο. Αυτές οι μεταβολές μπορούν να απομονωθούν και να προκύψει η αποεποχικοποιημένη χρονοσειρά, που δεν λαμβάνει υπόψη τις αλλαγές στα δεδομένα λόγω του χρονικού διαστήματος στο οποίο βρισκόμαστε.

- **Κυκλικότητα** είναι η συμπεριφορά που τακτικές διακυμάνσεις των δεδομένων δεν είναι προκαθορισμένες και ακολουθούν ένα συγκεκριμένο μοτίβο ανά διαφορετικά χρονικά διαστήματα. Η κυκλικότητα δεν πρέπει να συγχέεται με την εποχικότητα, καθώς η εποχικότητα ακολουθεί ένα συγκεκριμένο μοτίβο σε ένα γνωστό διάστημα. Οι περίοδοι κυκλικότητας συνήθως δεν είναι σταθερές και είναι μεγαλύτερες των δύο χρόνων. Ένα παράδειγμα κυκλικής συμπεριφοράς είναι η μείωση του πληθυσμού σε ένα οικοσύστημα, εξαιτίας της μείωσης των πηγών τροφής, και μόλις ο πληθυσμός φτάσει σε ένα χαμηλό σημείο, οι πηγές τροφής θα ανακάμψουν και ο πληθυσμός του οικοσυστήματος θα αρχίσει πάλι να αυξάνεται.
- **Ασυνέχεια (Disruption)** παρουσιάζεται στις χρονοσειρές εξαιτίας δεδομένων που δεν μπορούν να προβλεφθούν βάσει μοντέλων εποχικότητας, τάσης και κύκλου. Οι ασυνέχειες αποτελούν απότομες αλλαγές στην συμπεριφορά της χρονοσειράς και προκύπτουν από τυχαίους παράγοντες. Μία εταιρία που δεν μπορεί να παράξει προϊόντα εξαιτίας μίας απεργίας θα έχει πτώση στις πωλήσεις της δημιουργώντας μίας ασυνέχεια στα δεδομένα. Οι αλλαγές με παροδικό χαρακτήρα ονομάζονται ακραίες τιμές (outliers) ενώ αλλαγές συνήθως μόνιμου χαρακτήρα (μπορούν να υπάρξουν και αλλαγές που επηρεάζουν την χρονοσειρά για κάποιο χρονικό διάστημα και μετά επανέρχονται) που επηρεάζουν το επίπεδο της χρονοσειράς ονομάζονται αλλαγές επιπέδου (level-shifts). Ένα παράδειγμα αλλαγής επιπέδου ήταν η αρχή της πανδημίας του κορονοϊού, που προκάλεσε τεράστια πτώση της αγοράς.



Εικόνα 10: Ανάλυση χαρακτηριστικών χρονοσειράς

4.3 Μοντέλο χρονοσειρών

Τα μοντέλα χρονοσειρών αποτελούν μία από τις δημοφιλέστερες μεθόδους πρόβλεψης και βασίζονται σε ιστορικά δεδομένα στο πεδίο του χρόνου (ώρες, μέρες, μήνες, χρόνια) για την πρόβλεψη της μελλοντικής τιμής. Για την πρόβλεψη των μελλοντικών τιμών το μοντέλο χρονοσειράς τροφοδοτείται με τα ιστορικά δεδομένα, και η έξοδος που παράγει είναι η ζητούμενη πρόβλεψη. Η σχέση εισόδου και εξόδου είναι η εξής:

$$Y = f(X_1, X_2, \dots, X_n)$$

όπου τα X_1, X_2, \dots, X_n είναι τα ιστορικά δεδομένα με X_1 την παλαιότερη και X_n την πιο πρόσφατη τιμή.

Έτσι φαίνεται ότι ο κύριος στόχος του μοντέλου χρονοσειράς είναι η προσεκτική συλλογή και μελέτη των παρελθοντικών παρατηρήσεων για την ανάπτυξη κατάλληλου μοντέλου πρόβλεψης που περιγράφει καλύτερα την δεδομένη χρονοσειρά. Μέσω της ανάλυσης μπορούν να αντληθούν συμπεράσματα για τους εγγενείς παράγοντες της χρονοσειράς. Ένα τέτοιο παράδειγμα είναι η συνεχής παρατήρηση ίδιων τιμών στην χρονοσειρά εξαιτίας κυκλικότητας. Παρόλα αυτά τα μοντέλα χρονοσειράς δεν μπορούν να ενισχυθούν από πιθανώς χρήσιμες πληροφορίες, σε σχέση με τα αιτιοκρατικά, δίνοντας το πλεονέκτημα σε αυτά σε μερικές περιπτώσεις. Παρακάτω παρουσιάζονται κάποια από τα πιο κλασσικά μοντέλα χρονοσειρών.

4.3.1 Αφελής μέθοδος (Naive)

Η αφελής μέθοδος, αποτελεί το πιο απλοϊκό μοντέλο πρόβλεψης και χρησιμοποιείται μόνο σε δεδομένα χρονοσειρών. Για την παραγωγή της πρόβλεψης με την αφελή μέθοδο υποθέτουμε ότι η μελλοντική τιμή θα είναι ίδια με αμέσως προηγούμενη τιμή του μεγέθους, έτσι:

$$F(t) = Y(t - 1)$$

Όπου t είναι η χρονική περίοδος, $F(t)$ είναι η πρόβλεψη της μελλοντικής τιμής την περίοδο t και $Y(t-1)$ είναι η πραγματική τιμή της χρονοσειράς την αμέσως προηγούμενη χρονική περίοδο $t-1$.

Το σημαντικότερο πλεονέκτημα της μεθόδου είναι η απλοϊκή της μορφή και το χαμηλό υπολογιστικό κόστος. Παρόλα αυτά οι παραγόμενες προβλέψεις δεν είναι συνήθως ακριβείς και η συγκεκριμένη μέθοδος χρησιμοποιείται σαν σημείο αναφοράς και σύγκρισης για άλλα μοντέλα.

4.3.2 Εποχική αφελής μέθοδος (Seasonal Naive)

Η εποχική αφελής μέθοδος λαμβάνει υπόψη την εποχικότητα της χρονοσειράς, θέτοντας την τιμή της πρόβλεψης ίση με την τελευταία παρελθοντική τιμή της ίδιας εποχής. Για παράδειγμα, η πρόβλεψη της τιμής κάθε Απριλίου σε μία χρονοσειρά θα είναι η τιμή της χρονοσειράς του περασμένου Απριλίου. Έτσι το μοντέλο είναι το εξής:

$$F(t) = Y(t - m(k + 1))$$

Όπου m είναι η εποχική περίοδος και k είναι το ακέραιο μέρος που δείχνει τις εποχικές περιόδους που έχουν περάσει από το τελευταίο δεδομένο.

Η συγκεκριμένη μέθοδος χρησιμοποιείται σε δεδομένα με έντονο εποχιακό χαρακτήρα, αλλά και αυτή δεν παρέχει αρκετά ικανοποιητική ακρίβεια προβλέψεων.

4.3.3 Μέθοδος κινητού μέσου όρου (ΚΜΟ)

Ο μέσος όρος, στην στατιστική, είναι ο υπολογισμός της μέσης τιμής ενός συγκεκριμένου αριθμού δεδομένων. Η μέθοδος του κινητού μέσου όρου χρησιμοποιεί έναν σταθερό αριθμό (k) παρατηρήσεων, και υπολογίζει τον μέσο όρο για την παραγωγή της επόμενης πρόβλεψης. Έτσι το μοντέλο αυτό λαμβάνει υπόψη μόνο τα k πρόσφατα δεδομένα για την παραγωγή της πρόβλεψης. Μαθηματικά το μοντέλο διατυπώνεται ως εξής:

$$F(t + 1) = \frac{1}{k} * \sum_{i=t-k+1}^t Y(i)$$

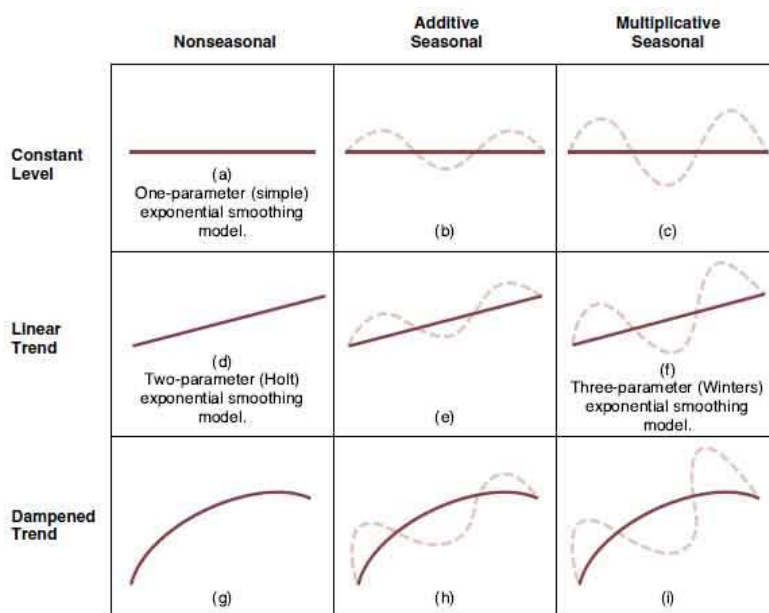
Στην περίπτωση που το k ισούται με την μονάδα το μοντέλο ΚΜΟ συμπίπτει με την Naïve μέθοδο. Ένα κύριο μειονέκτημα του ΚΜΟ είναι ότι επηρεάζεται από ακραίες τιμές, γιατί το μοντέλο χρησιμοποιεί μόνο τα τελευταία k δεδομένα και σε περιπτώσεις απότομων μεταβολών λόγω ασυνέχειας, οι προβλέψεις του μοντέλου τείνουν να ακολουθούν αυτές τις απότομες μεταβολές. Παράλληλα, επειδή η πρόβλεψη αποτελεί τον μέσο όρο των προηγούμενων παρατηρήσεων το μοντέλο δυσκολεύεται και στην πρόβλεψη απότομων αλλαγών επιπέδου και στην πρόβλεψη τάσης. Έτσι χρησιμοποιείται κυρίως σε στάσιμες χρονοσειρές ή για βραχυπρόθεσμες προβλέψεις.

4.3.4 Μέθοδος εκθετικής εξομάλυνσης

Η μέθοδος εκθετικής εξομάλυνσης είναι ένας τρόπος πρόβλεψης χρονοσειρών βάσει ενός σταθμισμένου αθροίσματος παρελθοντικών παρατηρήσεων. Η διαφορά με το μοντέλο του κινητού μέσου όρου είναι η χρήση φθινόντων βαρών στον υπολογισμό του μέσου όρου. Αυτός ο τρόπος εξομάλυνσης προσπαθεί να εξαλείψει τον θόρυβο που υπάρχει στην χρονοσειρά για τον καλύτερο υπολογισμό της πρόβλεψης. Η βασική ιδέα πίσω από τις μεθόδους εκθετικής εξομάλυνσης είναι ότι οι πρόσφατες τιμές επηρεάζουν περισσότερο τις προβλέψεις σε σχέση με τις παρελθοντικές τιμές. Έτσι τα μοντέλα τελικά υπολογίζουν την πρόβλεψη χρησιμοποιώντας εκθετικά μειούμενα βάρη από τα νεότερα στα παλαιότερα δεδομένα.

Η μέθοδος εξομάλυνσης χωρίζεται σε κατηγορίες αναλόγως του μοτίβου τάσης που ακολουθούν τα ιστορικά δεδομένα, ως προς το χρόνο:

- Το μοντέλο σταθερού επιπέδου (constant level), όπου η χρονοσειρά διατηρεί ένα σταθερό επίπεδο στην πάροδο του χρόνου.
- Το μοντέλο γραμμικής τάσης (linear trend) όπου το επίπεδο της χρονοσειράς μεταβάλλεται γραμμικά στην πάροδο του χρόνου και παρουσιάζει ανοδική ή καθοδική τάση.
- Το μοντέλο εκθετικής τάσης, όπου το επίπεδο της χρονοσειράς παρουσιάζει συνεχώς αυξανόμενη τάση. Το μοντέλο αυτό έχει αποδειχθεί ότι είναι υπεραισιόδοξο σε μακροχρόνιες προβλέψεις.
- Το μοντέλο φθίνουσας τάσης, όπου το επίπεδο της χρονοσειράς παρουσιάζει φθίνουσα μεταβαλλόμενη τάση.



Εικόνα 11: Μοντέλα εκθετικής εξομάλυνσης. Gardner Jr., E.S., *Exponential Smoothing Models* (1987). *International journal of forecasting*

4.3.4.1 Απλή εκθετική εξομάλυνση (Simple Exponential Smoothing - SES)

Το μοντέλο της απλής εκθετικής εξομάλυνσης είναι από τα δημοφιλέστερα μοντέλα εξομάλυνσης και βασίζεται στην απλή υπόθεση της προσθήκης φθινόντων βαρών στα παρελθοντικά δεδομένα. Έτσι για την πρόβλεψη την χρονική στιγμή t θα έχουμε:

$$\bar{S}_t = aS_t + a(1-a)S_{t-1} + a(1-a)^2S_{t-2} + \dots + a(1-a)^kS_{t-k} = F_t$$

Όπου το \bar{S}_t αποτελεί την πρόβλεψη του μοντέλου για την χρονική στιγμή t και αντιστοιχεί στο F_t που αναφέρθηκε στα παραπάνω μοντέλα και S_t είναι η πραγματική τιμή την στιγμή t σε αντιστοιχία με το Y_t .

Για τον γρήγορο υπολογισμό του μοντέλου παρατηρούμε ότι το άθροισμα των $k-1$ όρων είναι η πρόβλεψη την στιγμή $t-1$, δηλαδή:

$$aS_t + a(1-a)S_{t-1} + a(1-a)^2S_{t-2} + \dots + a(1-a)^kS_{t-k} = aS_t + (1-a)\bar{S}_{t-1} = \bar{S}_t$$

Έτσι η πρόβλεψη του μοντέλου θα βασίζεται σε ένα σταθμισμένο άθροισμα όλων των παλαιότερων προβλέψεων και θα είναι:

$$e_t = S_t - F_t$$

$$F_t = F_{t-1} + \alpha * e_t$$

Όπου από τις δύο αυτές σχέσεις προκύπτει το τελικό συμπυκνόμενο μοντέλο:

$$F_t = aS_t + (1-a)F_{t-1}$$

Στα παραπάνω το e_t είναι το σφάλμα της πρόβλεψης την χρονική στιγμή t , το S είναι το επίπεδο, F είναι η πρόβλεψη του μοντέλου την στιγμή t και α ονομάζεται η σταθερά εξομάλυνσης που παίρνει τιμές ανάμεσα στα $[0,1]$. Παρατηρούμε ότι η πρόβλεψη για την χρονική στιγμή t είναι το άθροισμα του σταθμισμένου μέσου όρου της τωρινής παρατήρησης Y και την προηγούμενης πρόβλεψης για $t-1$. Ο καθορισμός του

συντελεστή εξομάλυνσης α εξαρτάται από τον θόρυβο και την σταθερότητα του μέσου όρου της χρονοσειράς. Αν ο μέσος της χρονοσειράς υφίσταται μεγάλες αλλαγές τότε το α θα πρέπει να έχει μικρή τιμή για να είναι μικρή η επιρροή των πρόσφατων τιμών (Brown, 1956).

Καθώς το μοντέλο είναι αναδρομικό, είναι αναγκαία η εύρεση της αρχικής συνθήκης, δηλαδή του αρχικού επιπέδου $S(0)$, έτσι ώστε να ξεκινήσει η διαδικασία υπολογισμού του μοντέλου. Ως αρχικό επίπεδο, συνήθως χρησιμοποιείται:

- Ο μέσος όρος των παρατηρήσεων
- Το σταθερό επίπεδο από το μοντέλο της απλής γραμμικής παλινδρόμησης
- Η πρώτη παρατήρηση
- Το σταθερό επίπεδο από άλλο στατιστικό μοντέλο

4.3.4.2 Μοντέλο γραμμικής τάσης (Holt Exponential Smoothing)

Το μοντέλο γραμμικής τάσης αποτελεί εξέλιξη του μοντέλου απλής εκθετικής εξομάλυνσης που προτάθηκε από τον Robert Goodell Brown το 1956 (Brown, 1956) και επεκτάθηκε το 1957 από τον Charles C. Holt (Holt, 1957). Το μοντέλο της γραμμικής τάσης χωρίζεται σε δύο εξισώσεις, η μία του επιπέδου της χρονοσειράς και η άλλη της τάσης. Έτσι το μοντέλο επεκτείνει το μοντέλο της απλής εκθετικής εξομάλυνσης όπως φαίνεται παρακάτω:

$$e_t = S_t - F_t$$

$$T_t = T_{t-1} + \beta * e_t$$

$$F_{t+m} = F_t + m * e_t$$

$$F_t = T_{t-1} + F_{t-1} + \alpha * e_t$$

Και το τελικό μοντέλο είναι:

$$F_t = \alpha S_t + (1 - \alpha)F_{t-1} + T_{t-1}$$

$$T_t = T_{t-1} + \beta * (S_t - F_t)$$

Όπου το T_t δηλώνει την τάση, το m είναι ο χρονικός ορίζοντας της πρόβλεψης και το β είναι η σταθερά εξομάλυνσης της τάσης.

Στο συγκεκριμένο μοντέλο είναι αναγκαίο να προσδιοριστούν τόσο το αρχικό επίπεδο όσο και ο συντελεστής εξομάλυνσης της τάσης. Ο συντελεστής αυτός μπορεί να οριστεί ως την διαφορά της τελευταίας και της πρώτης παρατήρησης διαιρεμένης με το πλήθος των δειγμάτων ή ως η σταθερά κλίσης από το μοντέλο της απλής γραμμικής παλινδρόμησης.

Η μεγαλύτερη διαφορά αυτού του μοντέλου είναι ο συνυπολογισμός της τάσης κάνοντας πιο χρήσιμο το μοντέλο για μεσοπρόθεσμη πρόβλεψη δεδομένων. Τώρα τα δεδομένα μπορούν να έχουν ανοδική ή καθοδική τάση προσφέροντας μία γραμμική σχέση αύξησης ή μείωσης με τα πραγματικά δεδομένα.

4.3.4.3 Μοντέλο μη γραμμικής τάσης (Damped Exponential Smoothing)

Έχει παρατηρηθεί ότι το προτεινόμενο μοντέλο του Holt της γραμμικής τάσης, σε μακροχρόνιες περιόδους, δεν δουλεύει ικανοποιητικά και παράγει μοντέλα στα οποία

υπάρχει υπερεκτίμηση των τιμών (Gardner, 2015). Με αυτήν την λογική είχε προταθεί το 1985 από τους Gardner και McKenzie ένα μοντέλο που τροποποιούσε αυτό του Holt, με την προσθήκη ενός όρου για την εκθετική μείωση της τάσης με την πάροδο του χρόνου (Gardner & McKenzie, 1985). Το μοντέλο αυτό είναι το μοντέλο μη γραμμικής τάσης και έχει τις εξής εξισώσεις:

$$e_t = S_t - F_t$$

$$T_t = \varphi * T_{t-1} + \beta * e_t$$

$$F_t = F_{t-1} + \varphi * T_{t-1} + \alpha * e_t$$

$$F_{t+m} = F_t + \sum_{i=1}^m \varphi^i * e_t$$

Η καινούργια παράμετρος εδώ αποτελεί το φ , το οποίο εκφράζει την διόρθωση της τάσης. Ο υπολογισμός της τάσης όπως υποδεικνύει το μοντέλο δεν είναι γραμμικός. Ανάλογα με την τιμή που παίρνει η παράμετρος φ κάθε φορά έχουμε τα εξής:

- Με $\varphi = 0$ το μοντέλο εκφυλίζεται σε αυτό την απλής εκθετικής εξομάλυνσης, γιατί μηδενίζεται η τάση.
- Με $0 < \varphi < 1$ έχουμε το μοντέλο της μη γραμμικής τάσης που χρησιμοποιείται για μακροπρόθεσμες προβλέψεις.
- Με $\varphi = 1$ έχουμε το μοντέλο της γραμμικής τάσης αφού το S_t επηρεάζεται από το T_{t-1} χωρίς την παρουσία του φ , και το άθροισμα του σφάλματος αντικαθίσταται από το γινόμενο του m με την T_t .
- Με $\varphi > 1$ έχουμε το μοντέλο εκθετικής τάσης που έχει μεγάλη θετική προκατάληψη και τα αποτελέσματα είναι αρκετά υπεροπτικά.

Το συγκεκριμένο μοντέλο επεκτείνει το μοντέλο γραμμικής τάσης του Holt, κρατώντας την ακρίβεια για βραχυπρόθεσμες προβλέψεις, έχοντας όμως παράλληλα το πλεονέκτημα της βελτίωσης της ακρίβειας σε μακροπρόθεσμες προβλέψεις, κάτι που το μοντέλο της γραμμικής τάσης υστερεί (Gardner & McKenzie, 1985). Έτσι το μοντέλο χρησιμοποιείται συνήθως για μακροπρόθεσμες προβλέψεις γιατί προσφέρει μεγαλύτερη ακρίβεια σε σχέση με τα προηγούμενα αναφερθέντα μοντέλα. Από τα παραπάνω είναι εύκολα κατανοητό ότι η τιμή του φ που προτείνεται να επιλεγεί είναι ανάμεσα στα $[0, 1]$ για την χρήση του μοντέλου της μη γραμμικής τάσης. Για τον υπολογισμό των άλλων παραμέτρων, α και β , στοχεύουμε στην μείωση του τετραγωνικού σφάλματος και η επιλογή του αρχικού επιπέδου υπολογίζεται με ότι αναφέραμε παραπάνω.

4.3.5 Αυτοπαλινδρομικά μοντέλα κινητού μέσου όρου (ARIMA)

Τα μοντέλα ARIMA (autoregressive integrated moving averages) αποτελούν ακόμη μια δημοφιλή προσέγγιση στην πρόβλεψη χρονοσειρών, ίσως τη δημοφιλέστερη μαζί με την εκθετική εξομάλυνση. Μελετήθηκαν από τους (Box, et al., 2015) και είναι στοχαστικά μαθηματικά μοντέλα τα οποία βοηθάνε στην ανάλυση και πρόβλεψη της εξέλιξης μεγεθών. Σε αντίθεση με τα ντετερμινιστικά μοντέλα (βλ. γραμμική παλινδρόμηση), η χρήση των οποίων απαιτεί γνώση των παραγόντων από τις οποίες εξαρτάται το μέγεθος και όπου ο πλήρης εντοπισμός, μέτρηση και πρόβλεψή τους είναι

πρακτικά αδύνατος, η εφαρμογή των μοντέλων ARIMA βασίζεται στον υπολογισμό της πιθανότητας για την οποία η τιμή του μεγέθους βρίσκεται εντός κάποιου διαστήματος. Τα μοντέλα αυτά προτιμώνται επειδή παρουσιάζουν μεγάλη ακρίβεια σε βραχυπρόθεσμες προβλέψεις και δεν χρειάζονται κάποια σχετική επεξεργασία των δεδομένων όπως αφαίρεση τάσης ή εποχικότητας.

Ένα μοντέλο ARIMA αποτελείται από έναν παράγοντα AR, έναν παράγοντα MA και έναν παράγοντα διαφορίσης που εξασφαλίζει τη στασιμότητα της χρονοσειράς. Στις παρακάτω ενότητες περιγράφουμε τους επιμέρους παράγοντες και το συνολικό μοντέλο.

4.3.5.1 Μοντέλα AR

Το κομμάτι AR (Autoregressive) του ARIMA, υποδεικνύει ότι τα μοντέλα αποτελούνται αποκλειστικά από παράγοντες αυτοπαλινδρόμησης και συντομογραφικά γράφονται ως AR(p). Το p υποδηλώνει την τάξη των παραγόντων αυτοπαλινδρόμησης. Έτσι η πρόβλεψη υπολογίζεται ως το άθροισμα των τελευταίων p όρων και δίνεται από τον παρακάτω τύπο:

$$y_t = c + \varphi_1 * y_{t-1} + \varphi_2 * y_{t-2} + \dots + \varphi_p * y_{t-p} + Z_t$$

Όπου φ_i είναι οι συντελεστές αυτοσυσχέτισης του μοντέλου, c μία σταθερά και Z_t είναι λευκός θόρυβος. Στην ουσία, η τιμή της παρατήρησης y_t εξαρτάται κατά παράγοντα φ_1 από την προηγούμενη παρατήρηση, κατά παράγοντα φ_2 από την προ-προηγούμενη παρατήρηση κλπ, και έτσι υπολογίζεται ως γραμμικός συνδυασμός αυτών. Αν ορίσουμε ως τελεστή ολίσθησης το B ώστε:

$$By_t = y_{t-1} \text{ και } B^2 y_t = y_{t-2}$$

Τότε χρησιμοποιώντας αυτόν η σχέση μετατρέπεται στην:

$$(1 - \varphi_1 * B - \varphi_2 * B^2 - \dots - \varphi_p * B^p) * \bar{y}_t = Z_t$$

Οι ιδιότητες του μοντέλου μπορούν να βρεθούν μελετώντας τις ιδιότητες της συνάρτησης φ .

4.3.5.2 Μοντέλα MA

Ο παράγοντας MA (Moving average) του ARIMA, υποδεικνύει ότι το σφάλμα της πρόβλεψης στο παλινδρομικό μοντέλο είναι γραμμικός συνδυασμός των άλλων προηγούμενων σφαλμάτων. Έτσι τα μοντέλα αυτά είναι της τάξης q, όπου q είναι το πλήθος των παρελθοντικών σφαλμάτων που υπολογίζονται στον κινούμενο μέσο όρο του μοντέλου όπου το μοντέλο γράφεται ως MA(q). Η μαθηματική σχέση που προκύπτει είναι:

$$y_t = c + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

Όπου θ_i είναι οι συντελεστές μερικής αυτοσυσχέτισης του μοντέλου. Χρησιμοποιώντας τον τελεστή ολίσθησης η σχέση μετατρέπεται στην:

$$(1 - \varphi_1 * B - \varphi_2 * B^2 - \dots - \varphi_q * B^q) * \varepsilon_t = \bar{y}_t$$

4.3.5.3 Μοντέλα ARIMA

Τα μοντέλα AR(p) και MA(q) μπορούν να συνδυαστούν αποδοτικά για την ανάλυση και πρόβλεψη στάσιμων χρονοσειρών. Έτσι εισάγοντας στην εξίσωση και τον παράγοντα διαφορίσης προκύπτει ένα μεικτό μοντέλο το οποίο είναι και παλινδρομικό τάξης p και μέσου όρου τάξης q ονομάζεται ARIMA(p, q). Το ολοκληρωμένο παλινδρομικό μοντέλο δίνεται από την σχέση:

$$(1 - \varphi_1 * B_1 - \varphi_2 * B_2 - \dots - \varphi_p * B_p)(1 - B)^n(1 - B^m)^N * y_t = c + (1 - \theta_1 * B_1 - \theta_2 * B_2 - \dots - \theta_q * B_q) * \varepsilon_t$$

Ο πρώτος όρος του πρώτου μέλους δίνει τον παράγοντα AR, ο δεύτερος όρος δίνει τον παράγοντα διαφορίσης I(d) τάξης d, και ο όρος στο δεύτερο μέρος της εξίσωσης δίνει τον παράγοντα MA του μοντέλου. Έτσι στην γενική μορφή το ολοκληρωμένο μοντέλο τα μοντέλα ARIMA συντίθενται από έναν τυχαίο παράγοντα, από παρελθοντικές τιμές του μεγέθους που μελετάται και από σχετικούς στοχαστικούς παράγοντες. Στόχος μας είναι να βρούμε ποιος γραμμικός συνδυασμός αυτών των παραγόντων δίνει τα καλύτερα αποτελέσματα. Στην πράξη δεν μπορούμε ποτέ να είμαστε σίγουροι ποιος είναι ο βέλτιστος γραμμικός συνδυασμός, παρόλα αυτά μπορούμε να τον προσεγγίσουμε ικανοποιητικά.

Υποθέτοντας ότι το πολυώνυμο $(1 - \varphi_1 * B_1 - \varphi_2 * B_2 - \dots - \varphi_p * B_p)(1 - B)^n(1 - B^m)^N$ έχει ρίζες πολλαπλότητας d, το μοντέλο μπορεί να γραφεί ως ARIMA(p,d,q):

$$(1 - \varphi_1 * B_1 - \varphi_2 * B_2 - \dots - \varphi_p * B_p)(1 - B)^n(1 - B^m)^N = (1 - \theta_1 * B_1 - \theta_2 * B_2 - \dots - \theta_q * B_q) * (1 - B)^d$$

Η επιλογή μοντέλου ARIMA, ανάγεται στον προσδιορισμό των παραμέτρων p,q των τριών επιμέρους μοντέλων. Η εύρεση των κατάλληλων p και q μπορεί να γίνει αυτόματα μέσω του κριτηρίου πληροφορίας AIC.

4.4 Αιτιοκρατικό μοντέλο

Το αιτιοκρατικό μοντέλο βασίζεται στην εξήγηση της σχέσης αιτίου-αιτιατού μεταξύ δύο ή περισσότερων μεταβλητών, των εξαρτημένων, και των ανεξάρτητων (Granger & Newbold, 2014). Η παραπάνω σχέση μπορεί να διατυπωθεί υπό την εξής μορφή:

$$Y_t = f(X_{1(t-1)}, X_{2(t-1)}, \dots, X_{n(t-1)})$$

όπου τα $X_{1(t-1)}, X_{2(t-1)}, \dots, X_{n(t-1)}$ είναι οι ανεξάρτητες μεταβλητές την χρονική στιγμή t-1 και Y_t είναι η εξαρτημένη μεταβλητή την αμέσως επόμενη χρονική στιγμή.

Η σχέση αυτή υποδηλώνει πρακτικά ότι το Y εξαρτάται από n ανεξάρτητες μεταβλητές βάσει των οποίων γίνεται η πρόβλεψη. Για το αιτιοκρατικό μοντέλο, η εξαρτημένη και ανεξάρτητη μεταβλητή βασίζονται η μία στην άλλη, γιατί παρέχουν πληροφορίες η μία για την άλλη. Ένα παράδειγμα μίας τέτοιας σχέσης είναι η πώληση οικιακών συσκευών σε σχέση με τις άδειες κτισίματος καινούργιων σπιτιών, όπου η πρώτη είναι η εξαρτημένη μεταβλητή και η δεύτερη η ανεξάρτητη.

Η διαφορά των δύο μοντέλων είναι ότι το μοντέλο χρονοσειρών βασίζεται στην φύση των δεδομένων και στην συμπεριφορά τους βάσει τον χρόνο, ενώ το αιτιοκρατικό

μοντέλο βασίζεται στην συσχέτιση δύο ή περισσότερων μεταβλητών που η μία επηρεάζει την άλλη. Ένα μειονέκτημα του αιτιοκρατικού μοντέλου είναι η ευαισθησία του μοντέλου στις αλλαγές των σχέσεων μεταξύ των εξαρτημένων μεταβλητών, και δημιουργούν την ανάγκη της καλής συσχέτισης των μεγεθών. Παρόλα αυτά είναι αρκετά χρήσιμα μοντέλα γιατί παρέχουν την δυνατότητα πρόβλεψης για διάφορους συνδυασμούς ανεξάρτητων μεταβλητών, και έτσι μπορούν να βρουν σχέσεις τις οποίες δεν μπορούν να αιτιολογήσουν τα μοντέλα χρονοσειρών. Για την δημιουργία των προβλέψεων, τα μοντέλα πρέπει να προβλέψουν και τις μελλοντικές τιμές των ανεξάρτητων μεταβλητών, πράγμα που συνήθως είναι αρκετά δύσκολο. Τα πιο δημοφιλή αιτιοκρατικά μοντέλα αποτελούν οι μέθοδοι παλινδρόμησης.

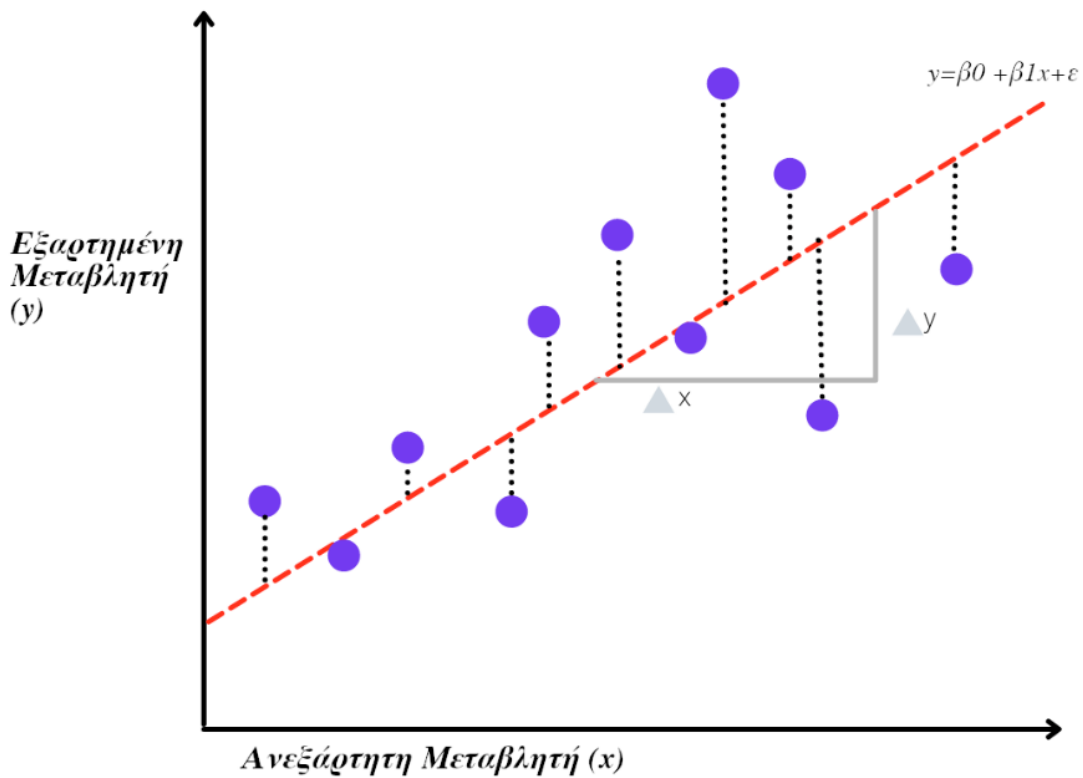
Οι μέθοδοι παλινδρόμησης βασίζονται στην λογική ότι υπάρχει συσχέτιση μεταξύ της χρονοσειράς y που προσπαθούμε να προβλέψουμε, με μία άλλη χρονοσειρά x . Οι μέθοδοι παλινδρόμησης χρησιμοποιούνται κυρίως για την εύρεση σχέσεων μεταξύ των δεδομένων, αλλά δευτερευόντως χρησιμοποιούνται και σαν βοηθητικό εργαλείο σε άλλες μεθόδους προβλέψεις (πχ. Στον υπολογισμό του σταθερού επιπέδου στο μοντέλο γραμμικής τάσης του Holt). Το βασικότερο μοντέλο είναι το μοντέλο της γραμμικής παλινδρόμησης που μελετάει την συσχέτιση μίας μεταβλητής x με την προβλεπόμενη μεταβλητή y , αλλά μπορεί να επεκταθεί με την χρήση παραπάνω από μίας μεταβλητής στο μοντέλο της πολλαπλής παλινδρόμησης. Τα δύο αυτά μοντέλα παρουσιάζονται παρακάτω.

4.4.1 Απλή γραμμική παλινδρόμηση

Η λογική της απλής γραμμικής παλινδρόμησης αποτελεί την ελαχιστοποίηση της απόστασης των παρατηρήσεων από την βέλτιστη γραμμή παλινδρόμησης. Η ευθεία υπολογίζεται με την ελαχιστοποίηση του σφάλματος και οι εξισώσεις που το μοντέλο χρησιμοποιεί στην απλούστερη μορφή του είναι:

$$y_t = \beta_0 + \beta_1 x_1 + \varepsilon_t$$

Όπου β_0 είναι η προβλεπόμενη τιμή όταν $x = 0$, β_1 είναι η κλίση της ευθείας και εκφράζει την μέση μεταβολή της y ανά μονάδα μεταβολής του x . Τέλος ε_t είναι το σφάλμα των παρατηρήσεων που πλησιάζει την ευθεία στις παρατηρήσεις και είναι η ποσότητα που το μοντέλο προσπαθεί να ελαχιστοποιήσει. Για την ελαχιστοποίηση χρησιμοποιείται η μέθοδος των ελαχίστων τετραγώνων όπου $\sum_{i=1}^n (Y_i - F_i)^2$ και επιδιώκεται η ελαχιστοποίηση αυτού του αθροίσματος. Παρόμοια με τα παραπάνω μοντέλα, με Y αναπαριστούμε την πραγματική τιμή την στιγμή i και με F την προβλεπόμενη τιμή από το μοντέλο την ίδια χρονική στιγμή.



Εικόνα 12: Γραφική αναπαράσταση ενός απλού μοντέλου γραμμικής παλινδρόμησης

Με την μέθοδο των ελαχίστων τετραγώνων ελαχιστοποιείται το σφάλμα και έτσι μπορούν να υπολογιστούν οι τιμές των β_1 και β_0 .

$$\min Error = \min \sum_{i=1}^n e^2 = \min \sum_{i=1}^n (Y_i - F_i)^2$$

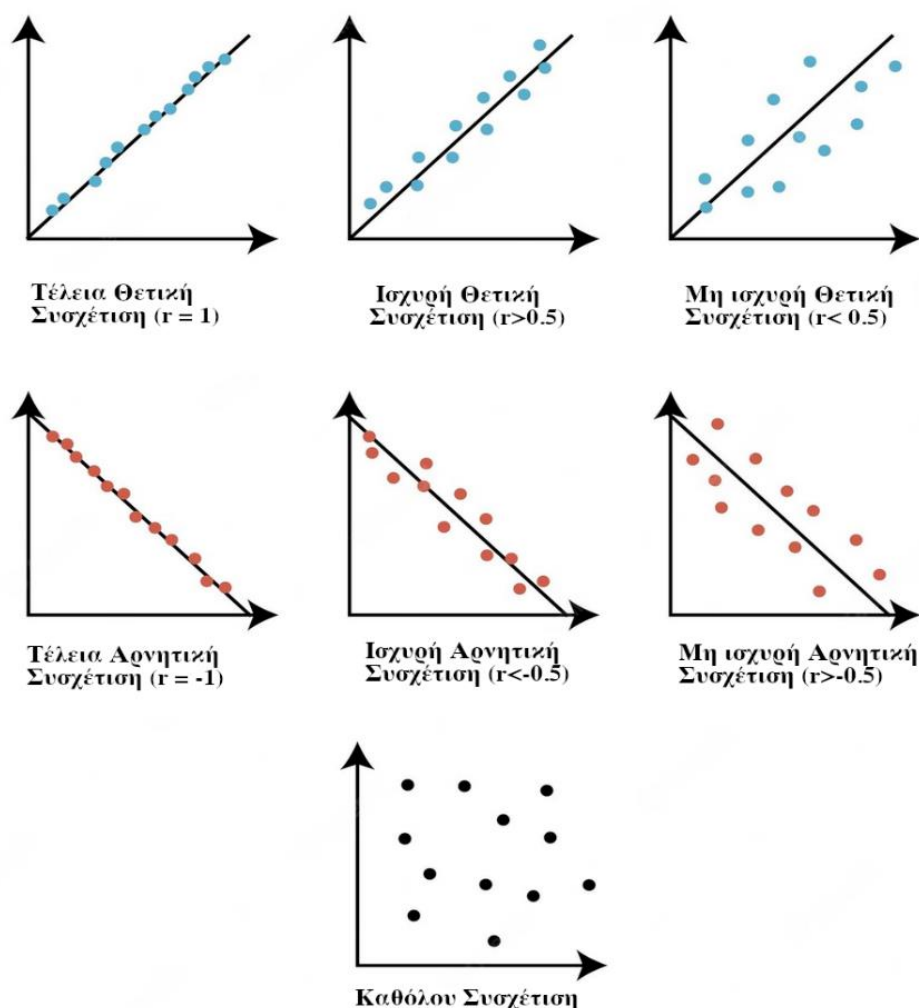
Και τελικά:

$$\begin{aligned} \beta_0 &= \bar{y} - \widehat{\beta}_1 \bar{x} \\ \beta_1 &= \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n (X_i Y_i) - \frac{1}{n} \sum_{i=1}^n X_i \sum_{j=1}^n Y_j}{\sum_{i=1}^n x_i^2 - \frac{1}{n} (\sum_{i=1}^n x_i)^2} = \frac{\overline{XY} - \bar{X}\bar{Y}}{\overline{X^2} - \bar{X}^2} \\ &= \frac{Cov[X, Y]}{Var[X, Y]} = r_{xy} \frac{s_y}{s_x} \end{aligned}$$

Όπου τα \bar{X} και \bar{Y} είναι οι μέσοι όροι των διακυμάνσεων x και y αντίστοιχα, r_{xy} ή απλά r είναι ένας συντελεστής που δείχνει την σχέση μεταξύ των x και y και ονομάζεται δειγματικός συντελεστής συσχέτισης ή συντελεστής Pearson και τα s_x και s_y είναι οι τυπικές αποκλίσεις των x και y . Τέλος τα $\widehat{\beta}_1$ και $\widehat{\beta}_0$ ονομάζονται εκτιμήτριες των ελαχίστων τετραγώνων και είναι πρακτικά οι τιμές των β_1 και β_0 που υπολογίζει το παραπάνω μοντέλο.

4.4.2 Δειγματικός συντελεστής συσχέτισης (Pearson)

Ο συντελεστής Pearson παίρνει τιμές στο διάστημα $[-1,1]$. Όσο πιο κοντά στο 1 είναι οι τιμές τόσο μεγαλύτερη είναι και η συσχέτιση μεταξύ των x και y , ενώ όσο πιο κοντά στο -1 , τόσο μεγαλύτερη αρνητική συσχέτιση έχουν (πρακτικά σημαίνει ότι όσο αυξάνεται το ένα μέγεθος μειώνεται το άλλο). Τέλος όσο πιο κοντά πλησιάζει στο 0 ο συντελεστής, τα μεγέθη είναι ασυσχέτιστα μεταξύ τους και δεν παρουσιάζουν κάποια κοινή συμπεριφορά.



Εικόνα 13: Παραδείγματα συντελεστή συσχέτισης Pearson

Καλύτερη φυσική ερμηνεία της συσχέτισης δύο μεταβλητών επιτυγχάνεται με τον συντελεστή r^2 που ονομάζεται συντελεστής προσδιορισμού. Ο συντελεστής προσδιορισμού δίνει το ποσοστό μεταβλητότητας των τιμών της Y που υπολογίζεται από τη X (και αντίστροφα) και είναι ένας χρήσιμος τρόπος να συνοψίσουμε τη σχέση δύο μεταβλητών. Παίρνει τιμές μεταξύ των $[0,1]$.

4.4.3 Σχέση του συντελεστή συσχέτισης και παλινδρόμησης

Για το μοντέλο της απλής γραμμικής παλινδρόμησης τα κάθε ζεύγος τιμών (β_1, β_0) καθορίζει μία διαφορετική γραμμική σχέση που εκφράζεται γεωμετρικά από μία ευθεία

γραμμή. Ο συντελεστής β_0 είναι η τιμή του y για $x = 0$ και λέγεται διαφορά ύψους (intercept), ενώ ο συντελεστής του x , β_1 είναι η κλίση (slope) της ευθείας ή αλλιώς συντελεστής παλινδρόμησης (regression coefficient).

Η σχέση μεταξύ του r και του $\widehat{\beta}_1$ (εκτίμηση του συντελεστή παλινδρόμησης) δίνεται ως εξής:

$$r = \widehat{\beta}_1 \frac{s_x}{s_y} \text{ ή } \widehat{\beta}_1 = r \frac{s_y}{s_x}$$

Και τα δύο μεγέθη εκφράζουν ποσοτικά τη γραμμική συσχέτιση των μεταβλητών x και y και η σχέση τους είναι η εξής:

- Αν η συσχέτιση είναι θετική ($r > 0$) τότε η κλίση της ευθείας παλινδρόμησης β_1 είναι επίσης θετική
- Αν η συσχέτιση είναι αρνητική ($r < 0$) τότε η κλίση της ευθείας παλινδρόμησης β_1 είναι επίσης αρνητική
- Αν οι μεταβλητές x και y δεν συσχετίζονται ($r = 0$) τότε η ευθεία παλινδρόμησης είναι οριζόντια ($\beta_1 = 0$)

4.4.4 Στατιστικοί δείκτες

Σε πολλά προβλήματα μπορεί οι μεταβλητές που μελετώνται να μην συσχετίζονται μεταξύ τους. Έτσι το πρόβλημα της απλής γραμμικής παλινδρόμησης μπορεί να εκφυλιστεί στο μοντέλο $y_t = \beta_0 + \varepsilon_t$, όπου η γραμμή παλινδρόμησης είναι οριζόντια και η μεταβλητή y_t δεν εξαρτάται από την μεταβλητή x . Το μοντέλο που προκύπτει είναι το μοντέλο της απλής γραμμικής παλινδρόμησης για $\beta_1 = 0$. Αυτή η υπόθεση και αναζήτηση του μοντέλου με τα καλύτερα αποτελέσματα, ονομάζεται μηδενική υπόθεση $H_0: \beta_1 = 0$. Αντίστοιχα μπορεί να γίνει και η υπόθεση της διαφοράς ύψους $H_0: \beta_0 = 0$.

Οι παραπάνω υποθέσεις για τους συντελεστές β_0 και β_1 γεννούν την ανάγκη για δημιουργία κάποιων δεικτών που ελέγχουν την σχέση των συντελεστών και πως αυτοί επηρεάζουν το μοντέλο της γραμμικής παλινδρόμησης. Έτσι έχουν δημιουργηθεί οι στατιστικοί δείκτες F και t , με τους οποίους υπολογίζεται η ακρίβεια των συντελεστών β_0 και β_1 και η αξιοπιστία του υπολογισμού της ευθείας παλινδρόμησης.

Ο πρώτος δείκτης που ορίζεται ονομάζεται στατιστικός δείκτης F και επιτρέπει την εκτίμηση της σημαντικότητας της εξίσωσης παλινδρόμησης, δηλαδή απαντάει στο ερώτημα του αν υπάρχει σχέση ανάμεσα στις εξαρτώμενες μεταβλητές.

Έτσι για τον δείκτη F μελετάμε δύο περιπτώσεις όπου:

1. Εκτιμάται το κανονικό (ολοκληρωμένο) μοντέλο, όπου βρίσκουμε τους εκτιμητές $\widehat{\beta}_0$, $\widehat{\beta}_1$ και υπολογίζουμε το τετραγωνικό σφάλμα των προβλέψεων το οποίο ορίζουμε ως $SSE(F) = \sum (Y_i - F_i)^2$.
2. Εκτιμάται το μοντέλο της μηδενικής υπόθεσης, όπου βρίσκουμε τον εκτιμητή $\widehat{\beta}_0$ και υπολογίζουμε το τετραγωνικό σφάλμα των προβλέψεων το οποίο ορίζουμε ως $SSE(R) = \sum (Y_i - \bar{Y})^2$.

Και τελικά ο δείκτης F είναι:

$$F = \frac{\frac{SSE(R) - SSE(F)}{df_R - df_F}}{\frac{SSE(F)}{df_F}} = \frac{\frac{\sum(F_i - \bar{Y})^2}{df_R - df_F}}{\frac{\sum(Y_i - F_i)^2}{df_F}}$$

Το τετραγωνικό σφάλμα του μοντέλου μηδενικής υπόθεσης SSE(R) δεν μπορεί ποτέ να είναι μικρότερο από το τετραγωνικό σφάλμα του ολοκληρωμένου μοντέλου SSE(F).

- Αν το SSE(F) είναι σχετικά κοντά με το SSE(R), τότε η διακύμανση του ολοκληρωμένου μοντέλου είναι περίπου η ίδια με αυτή του απλούστερου μοντέλου, και για χάρη υπολογισμών προτιμάται το απλοποιημένο μοντέλο.
- Από την άλλη όταν η διαφορά των σφαλμάτων είναι σημαντική το ολοκληρωμένο μοντέλο δίνει καλύτερα αποτελέσματα και προτιμάται.

Πρακτικά ο δείκτης F είναι η διαφορά του σφάλματος των δύο μοντέλων SSE(F) και SSE(R) και από τα παραπάνω προκύπτει ότι για μεγάλες τιμές του F προτιμάται το ολοκληρωμένο μοντέλο, ενώ για μικρές τιμές προτιμάται το απλοποιημένο μοντέλο. Τα df_F και df_R ονομάζονται βαθμοί ελευθερίας και σχετίζονται με τον αριθμό μεταβλητών του κάθε μοντέλου. Για το μοντέλο της απλής γραμμικής παλινδρόμησης θα έχουμε:

$$df_R = n - 1, df_F = n - 2$$

Παράλληλα, υπάρχουν και ο στατιστικός δείκτης t, που επιτρέπει την εκτίμηση της σημαντικότητας των συντελεστών του μοντέλου. Έτσι έχουμε τα τυπικά σφάλματα για τους δύο συντελεστές:

$$SE_{\beta_0} = \sqrt{\frac{\sum(Y_i - F_i)^2}{n - k}} * \sqrt{\frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^n (X_i - \bar{X})^2}}$$

$$SE_{\beta_1} = \sqrt{\frac{\sum(Y_i - F_i)^2}{n - k}} * \sqrt{\frac{1}{\sum_{i=1}^n (X_i - \bar{X})^2}}$$

Και οι δείκτες t υπολογίζονται ως:

$$t_{\beta_0} = \frac{\widehat{\beta}_0}{SE_{\beta_0}}, t_{\beta_1} = \frac{\widehat{\beta}_1}{SE_{\beta_1}}$$

Πρακτικά ο δείκτης F εξετάζει την σημαντικότητα του ίδιου του μοντέλου, ενώ ο δείκτης t εξετάζει την σημαντικότητα καθενός από τους συντελεστές της γραμμικής παλινδρόμησης. Όταν η τιμή του δείκτη t πλησιάζει στο 0, τόσο λιγότερο συνεισφέρει ο συγκεκριμένος συντελεστής στον υπολογισμό του μοντέλου. Η σχέση αυτή φαίνεται και στον παρακάτω πίνακα.

Tail Probability p												
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	.681	.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	.679	.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	.678	.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	.677	.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	.675	.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
∞	.674	.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%

Confidence level C

Εικόνα 14. Πίνακας εμπιστοσύνης t -values. t -Distribution Critical Values (2023). Woodstown.org

Ο πίνακας δίνει την πιθανότητα συνεισφοράς της ανεξάρτητης μεταβλητής σε σχέση με τους βαθμούς ελευθερίας και την τιμή του αντίστοιχου t -value. Η πρώτη γραμμή δίνει τον συντελεστή α (δίνει την κατανομή των t -values) και η πρώτη στήλη τον βαθμό ελευθερίας ($df = n - k$). Το C από την άλλη, δίνει την πιθανότητα να μην συνεισφέρει η αντίστοιχη μεταβλητή στο μοντέλο γραμμικής παλινδρόμησης. Από τον πίνακα είναι εμφανές ότι όταν το α , άρα και τα t -values, τείνει στο μηδέν, τόσο αυξάνεται και η πιθανότητα να μην συνεισφέρει η μεταβλητή στο μοντέλο.

4.4.5 Πολλαπλή γραμμική παλινδρόμηση

Το μοντέλο της απλής γραμμικής παλινδρόμησης μπορεί να επεκταθεί όταν υπάρχουν περισσότερες από μία ανεξάρτητες μεταβλητές X_i και ονομάζεται μοντέλο πολλαπλής γραμμικής παλινδρόμησης. Η γενική μορφή του μοντέλου είναι:

$$y_t = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon_t$$

Όπου συνολικά έχουμε k ανεξάρτητες μεταβλητές. Οι συντελεστές β_k δείχνουν την σημαντικότητα κάθε μεταβλητής στο μοντέλο.

Για τον υπολογισμό των συντελεστών του μοντέλου ακολουθούμε πάλι την μέθοδο ελαχίστων τετραγώνων και υπολογίζουμε το εξής άθροισμα:

$$\min \sum_{i=1}^n e^2 = \min \sum_{i=1}^n (Y_i - F_i)^2 = \min \sum_{i=1}^n (Y_i - b_0 - b_1 X_{1,i} - b_2 X_{2,i} + \dots)^2$$

Προκειμένου να βρούμε τις εκτιμήτριες των συντελεστών αρκεί να υπολογίσουμε τις μερικές παραγώγους για κάθε έναν από τους συντελεστές, να θέσουμε τις υπολογισμένες παραγώγους ίσες με μηδέν και να λύσουμε ένα γραμμικό σύστημα n εξισώσεων με n αγνώστους.

4.5 Δείκτες Ακρίβειας

Η μεγαλύτερη ανάγκη για την παραγωγή καλών προβλέψεων είναι η δυνατότητα σύγκρισης και αξιολόγησης όλων των μοντέλων σε σχέση με την ακρίβεια και αμεροληψία τους. Οι δείκτες σφάλματος/ακρίβειας βασίζονται στην σύγκριση την πραγματικής σε σχέση με την προβλεπόμενη τιμή για να βγάλουν τα αποτελέσματα τους. Έτσι ειδικά στις μεθόδους χρονοσειρών χρησιμοποιούνται διάφορες τέτοιες μέθοδοι για την αξιολόγηση των μοντέλων πρόβλεψης. Οι βασικότεροι δείκτες παρουσιάζονται παρακάτω.

4.5.1 Σφάλματα εξαρτώμενα από την κλίμακα

Το πρώτο είδος σφαλμάτων εξαρτάται στην υπολογισμό της τιμής του σφάλματος e_i το οποίο είναι:

$$e_t = y_t - f_t$$

Τα σφάλματα αυτά εκφράζονται στις μονάδες μέτρησης που εκφράζονται και οι παρατηρήσεις και το χαρακτηριστικό τους είναι ότι εξαρτώνται από την στάθμη της εκάστοτε χρονοσειράς. Αυτού του είδους τα σφάλματα έχουν φυσική σημασία που προκύπτει από την μονάδα μέτρησης. Επίσης μετράνε την προκατάληψη του μοντέλου, σε αντίθεση με τα ποσοστιαία που θα μελετηθούν παρακάτω. Από την άλλη αδυνατούν να δώσουν συμπεράσματα για την ακρίβεια του μοντέλου.

4.5.1.1 Μέσο σφάλμα (Mean Error - ME)

Το μέσο σφάλμα υπολογίζεται ως η μέση τιμή των παρατηρήσεων σε σχέση με τις προβλέψεις του μοντέλου και μαθηματικά δίνεται από την σχέση:

$$ME = \frac{1}{n} \sum_{t=1}^n Y_t - F_t$$

Όσο ο δείκτης τείνει στο μηδέν τόσο μικρότερα είναι τα σφάλματα. Θετικές τιμές του δείκτη δηλώνουν απαισιοδοξία στις προβλέψεις ενώ αρνητικές τιμές δηλώνουν το αντίθετο. Για όλα τα παραπάνω ο δείκτης αναφέρεται και ως bias (προκατάληψη) του μοντέλου.

4.5.1.2 Μέσο απόλυτο σφάλμα (Mean Absolute Error – MAE)

Το μέσο απόλυτο σφάλμα υπολογίζεται ως το άθροισμα της απόλυτης τιμής του σφάλματος των παρατηρήσεων και δίνεται από την σχέση:

$$MAE = \frac{1}{n} \sum_{t=1}^n |Y_t - F_t|$$

Ο συγκεκριμένος δείκτης δεν νοιάζεται για την αρνητική ή την θετική κατεύθυνση των προβλέψεων και δεν μπορεί να μας πει το αν το μοντέλο έχει προκατάληψη. Το μόνο που μας δείχνει είναι την ακρίβεια του μοντέλου που εφαρμόστηκε όσο μικραίνει η τιμή του σφάλματος.

4.5.1.3 Μέσο τετραγωνικό σφάλμα (Mean Squared Error – MSE)

Το μέσο τετραγωνικό σφάλμα είναι το άθροισμα των τετραγωνικών σφαλμάτων των παρατηρήσεων και δίνεται από την σχέση:

$$MSE = \frac{1}{n} \sum_{t=1}^n (Y_t - F_t)^2$$

Ο συγκεκριμένος δείκτης χρησιμοποιείται για τον υπολογισμό του ρίσκου και επίσης χρησιμοποιείται για τον υπολογισμό των βέλτιστων παραμέτρων εξομάλυνσης στην προσαρμογή των μοντέλων πρόβλεψης. Παρόλα αυτά λόγω του τετραγώνου, μεμονωμένα σφάλματα αυξάνουν περισσότερο τις τιμές των δεικτών, κάνοντας ευκολότερη την διάκριση ενός μοντέλου που παρέχει πιο ακριβείς προβλέψεις και παρόλο που δεν είναι απόλυτος, πάλι αδυνατεί να μετρήσει την προκατάληψη του μοντέλου.

4.5.1.4 Ρίζα μέσου τετραγωνικού σφάλματος (Root Mean Squared Error – RMSE)

Το μοντέλο της ρίζας του μέσου τετραγωνικού σφάλματος χρησιμοποιείται για τον υπολογισμό του μέσου όρου της διαφοράς της πραγματικής από την προβλεπόμενη τιμή και μαθηματικά δίνεται από την σχέση:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - F_t)^2}$$

Ο δείκτης αυτός δεν θα πρέπει να συγχέεται με τον δείκτη MAE, καθώς οι δύο δίνουν διαφορετικά αποτελέσματα. Ο δείκτης RMSE είναι η τετραγωνική διαφορά των σφαλμάτων και έτσι θα αυξηθεί περισσότερο σε τιμή με προβλέψεις που απέχουν πολύ από την πραγματική τιμή, όταν δηλαδή το σφάλμα αυξάνεται. Τα τετραγωνικά σφάλματα (RMSE και MSE) προτιμώνται όταν θέλουμε να δούμε την γενική ακρίβεια του μοντέλου σε σχέση με το πόσο κοντινές είναι οι προβλέψεις του. Παράλληλα, λόγω

του τετραγώνου, ο δείκτης επηρεάζεται περισσότερο από μεγαλύτερα σφάλματα όπως και ο δείκτης MSE, ενώ η τιμή του δείκτη MAE αυξάνεται γενικά με το πλήθος των σφαλμάτων και δίνει λιγότερη σημασία στην σχετική ακρίβεια των τιμών. Η διαφορά του RMSE και του MSE είναι ότι ο δείκτης RMSE εξαιτίας της ρίζας προσπαθεί να κανονικοποιήσει το σφάλμα στο επίπεδο του MAE για σύγκριση των δύο αυτών δεικτών.

4.5.2 Ποσοστιαία σφάλματα

Τα ποσοστιαία σφάλματα υπολογίζονται με βάση την τιμή P_i με την σχέση:

$$P_t = \frac{|e_t|}{y_t}$$

Αυτού του είδους τα σφάλματα είναι αυτά που συναντώνται πιο συχνά στον τομέα των προβλέψεων. Οι δείκτες των σφαλμάτων αυτών δεν δίνουν τιμές ανάλογες των προβλέψεων αλλά ένα ποσοστό σφάλματος, κάνοντας πιο εύκολη την σύγκριση των δεικτών μεταξύ τους με αποτέλεσμα να είναι προτιμότερη η χρήση τους για την σύγκριση διαφορετικών χρονοσειρών. Αυτός είναι και ο λόγος που συναντώνται συχνά στους τομείς των προβλέψεων.

4.5.2.1 Μέσο απόλυτο ποσοστιαίο σφάλμα (Mean Absolute Percentage Error – MAPE)

Το μέσο απόλυτο ποσοστιαίο σφάλμα εκφράζει, όπως αναφέρθηκε παραπάνω, τον δείκτη σφάλματος σαν ποσοστό. Η μαθηματική του σχέση είναι η εξής:

$$MAPE = \frac{100}{n} \sum_{t=1}^n \left| \frac{Y_t - F_t}{Y_t} \right|$$

Το σφάλμα της πρόβλεψης διαιρείται με την πραγματική τιμή και το άθροισμα των ποσοστιαίων σφαλμάτων αθροίζεται και διαιρείται με βάση το πλήθος των παρατηρήσεων για την παραγωγή του μέσου όρου. Έτσι έχουμε ένα ποσοστό επί τις εκατό που χρησιμοποιείται στην σύγκριση διάφορων μοντέλων.

Παρόλα αυτά ο δείκτης έχει κάποια μειονεκτήματα. Ένα από αυτά είναι ότι οδηγεί σε απροσδιοριστίες για μηδενικές και ακραίες πραγματικές τιμές εξαιτίας του παρονομαστή. Επίσης σε συνέχεια αυτού το διάστημα του δείκτη MAPE δεν είναι $[0, 100]$, αλλά από το μηδέν μέχρι το άπειρο, κάνοντας τον δείκτη πιο ασταθή σε ακραίες τιμές. Ένα δεύτερο μειονέκτημα είναι η υπερεκτίμηση σφαλμάτων με μικρές πραγματικές τιμές. Αν το Y έχει μικρή τιμή το ποσοστιαίο σφάλμα, μιας και ποσοστό της πραγματικής τιμής, είναι σημαντικά μεγάλο και επηρεάζει παραπάνω τον δείκτη.

4.5.2.2 Συμμετρικό μέσο απόλυτο ποσοστιαίο σφάλμα (Symmetric Mean Absolute Percentage Error – sMAPE)

Το συμμετρικό μέσο απόλυτο ποσοστιαίο σφάλμα αποτελεί παραλλαγή του MAPE όπου ο δείκτης τώρα διαιρείται με το ημι-άθροισμα της παρατήρησης και της πρόβλεψης. Η μαθηματική σχέση είναι:

$$sMAPE = \frac{100}{n} \sum_{t=1}^n \left| \frac{Y_t - F_t}{\frac{Y_t + F_t}{2}} \right|$$

Η διαφορά του sMAPE είναι ότι ο δείκτης αυτός αποκτά πάνω όριο και πλέον μπορεί να πάρει τιμές στο διάστημα [0%, 200%]. Με αυτό τον τρόπο, ξεπερνιούνται τα προβλήματα απροσδιοριστίας και ακραίων τιμών που προκύπτουν κατά τη χρήση του MAPE. Παράλληλα δεν έχουμε υπερεκτίμηση σφαλμάτων για τους ίδιους λόγους.

Παρόλα αυτά ένα σημαντικό πρόβλημα του sMAPE είναι πως σε περίπτωση που η πραγματική τιμή και η πρόβλεψη ισούνται με μηδέν, τότε το ποσοστό σφάλματος μεγιστοποιείται και κάνει τον δείκτη δύσκολο σε χρήση για μοντέλα με τιμές και προβλέψεις πολύ κοντά στο μηδέν.

4.5.3 Σχετικά σφάλματα

Τα σχετικά σφάλματα αποτελούν την κανονικοποίηση των προηγούμενων μεθόδων σφαλμάτων, χρησιμοποιώντας κάποιο βασικό μοντέλο το οποίο θεωρείται benchmark. Συνήθως για benchmark θεωρούμε την Naive μέθοδο, όπως αναφέρθηκε και παραπάνω. Έτσι το κανονικοποιημένο σφάλμα θα παίρνει την μορφή:

$$RelError = \frac{Error_{method}}{Error_{benchmark}}$$

Το αποτέλεσμα που προκύπτει μας δίνει να καταλάβουμε πόσο καλύτερη είναι η μέθοδος που χρησιμοποιούμε και αν έχει νόημα να θυσιαστεί η υπολογιστική ισχύς που χρειάζεται για την παραγωγή της πρόβλεψης.

4.5.3.1 Μέσο απόλυτο κανονικοποιημένο σφάλμα (Mean Absolute Scaled Error - MASE)

Το μέσο απόλυτο κανονικοποιημένο σφάλμα δίνει την ίδια βαρύτητα στα μικρά και μεγάλα σφάλματα σε αντίθεση με τους MSE και RMSE. Η μαθηματική του έκφραση είναι:

$$MASE = \frac{\frac{1}{h} \sum_{t=n+1}^{n+h} |Y_t - F_t|}{\frac{1}{n-1} \sum_{t=2}^n |Y_t - Y_{t-1}|}$$

Αποτελεί υποπερίπτωση σχετικού σφάλματος κατά την οποία κανονικοποιείται το σφάλμα της χρησιμοποιούμενης μεθόδου με τις προβλέψεις της Naive μεθόδου. Στην περίπτωση που το $MASE > 1$, τότε το προβλεπτικό μοντέλο έχει μεγαλύτερο σφάλμα από την Naive μέθοδο και πρακτικά σημαίνει ότι η μέθοδος πρόβλεψης δεν είναι αρκετά ικανοποιητική. Έτσι η συσχέτιση αυτή μπορεί να μας δώσει ένα άμεσο αποτέλεσμα για την ποιότητα των προβλέψεων.

Από την άλλη, η κατανόηση και εξήγηση του συγκεκριμένου δείκτη είναι δύσκολη. Η Naive μέθοδος χρησιμοποιείται στο σύνολο δεδομένων εκπαίδευσης (in-sample dataset), ενώ το μοντέλο κάνει τις προβλέψεις στο σύνολο δεδομένων αξιολόγησης (out-of-sample dataset). Έτσι η συσχέτιση μεταξύ των δύο είναι προβληματική καθώς αναφέρονται σε διαφορετικά σύνολα δεδομένων.

4.5.3.2 Ρίζα μέσου τετραγωνικού κανονικοποιημένου σφάλματος (Root Mean Squared Scaled Error - RMSSE)

Η ρίζα μέσου τετραγωνικού κανονικοποιημένου σφάλματος είναι ο δείκτης RMSE κανονικοποιημένος κατά τον τετραγωνικό παράγοντα πρόβλεψης της Naive. Έχει

παρόμοια χαρακτηριστικά με τον δείκτη MASE, με διαφορά το αυξημένο βάρος που δίνεται στα μεγάλα σφάλματα λόγω της τετραγωνικής διαφοράς που χρησιμοποιείται. Ο μαθηματικός τύπος είναι:

$$RMSSE = \sqrt{\frac{\frac{1}{n} \sum_{t=1}^n (Y_t - F_t)^2}{\frac{1}{n-1} \sum_{t=2}^n (Y_t - Y_{t-1})^2}}$$

Ο δείκτης αυτός, όπως και ο MASE είναι δείκτες σφαλμάτων με ορισμένα χρήσιμα χαρακτηριστικά. Πιο συγκεκριμένα, είναι ανεξάρτητοι της κλίμακας (ή στάθμης) των δεδομένων, δεν έχουν απροσδιοριστίες σε μηδενικά δεδομένα και αντιμετωπίζουν με συμμετρικό τρόπο τα αρνητικά και θετικά σφάλματα.

Κεφάλαιο 5. Πειραματική διαδικασία

Όπως αναφέρθηκε και πριν, τα κρυπτονομίσματα χρησιμοποιούνται ως επενδυτική επιλογή τόσο από τα χρηματοπιστωτικά ιδρύματα όσο και από μικρότερους επενδυτές. Επομένως μία καλή πρόβλεψη της τιμής αυτών είναι ένα πολύ σημαντικό εργαλείο στην διαδικασία λήψης ορθών αποφάσεων για την επένδυση ή όχι σε αυτά. Αν συμπεριληφθεί παράλληλα και ο ασταθής τους χαρακτήρας, που κάνει δύσκολη την λήψη εμπειρικών αποφάσεων, ένα στατιστικό εργαλείο πρόβλεψης παίζει καθοριστικό ρόλο. Στόχος του κεφαλαίου, λοιπόν, είναι η δημιουργία ενός μοντέλου συσχέτισης των διαφόρων τεχνολογιών των κρυπτονομισμάτων (χωρίζονται σε κατηγορίες βάσει του αλγορίθμου ομοφωνίας) και της εξαγωγής συμπερασμάτων που βοηθούν στην πρόβλεψη της τιμής αυτών. Ο κύριος πυλώνας του μοντέλου είναι η δημιουργία μίας μέσης τιμής για την κάθε τεχνολογία κρυπτονομισμάτων και η χρήση αυτών για την πρόβλεψη της τιμής των άλλων τεχνολογιών, και κατ' επέκταση μεμονωμένων κρυπτονομισμάτων, με χρήση του μοντέλου γραμμικής παλινδρόμησης.

5.1 Σύλλογή δεδομένων

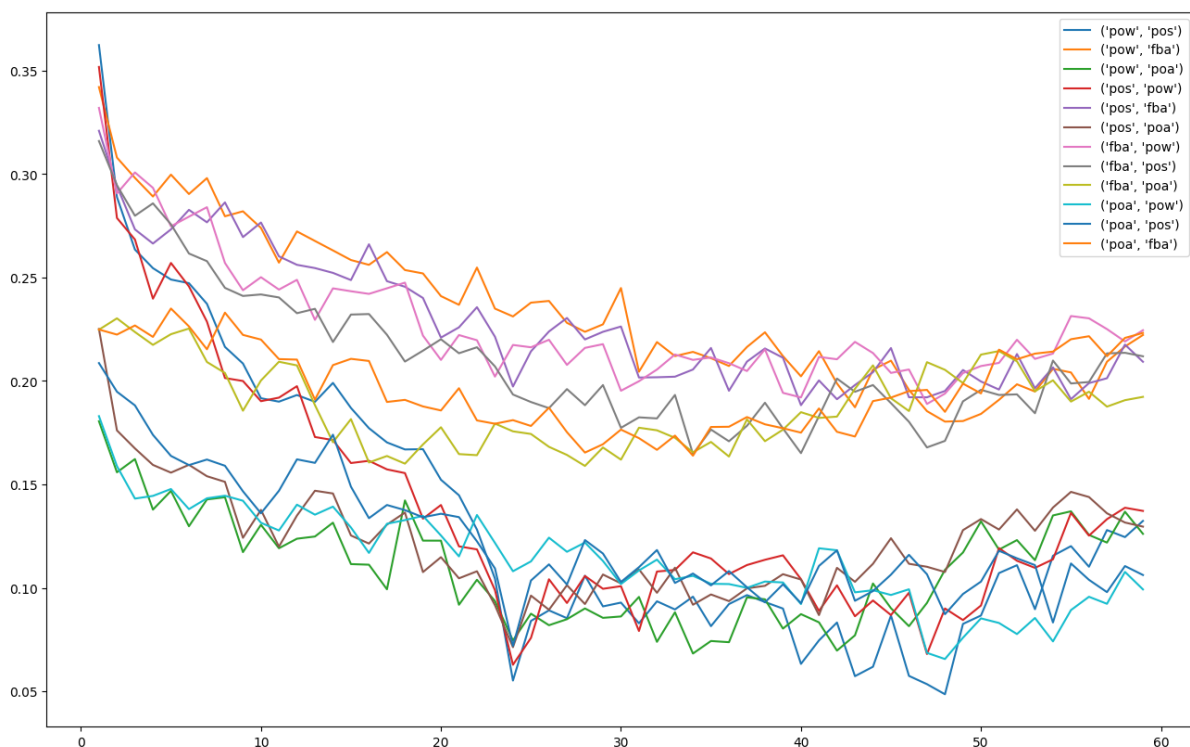
Ελλιπή ή κακής ποιότητας δεδομένα, που τροφοδοτούν ένα καλό προγραμματιστικό μοντέλο, τις περισσότερες φορές οδηγούν σε χειρότερα αποτελέσματα απ' ότι ένα απλό μοντέλο με καλής ποιότητας δεδομένα. Έτσι στόχος της μελέτης είναι η χρήση όσο το δυνατόν περισσότερων και καλύτερων ιστορικών δεδομένων για την σωστή εκπαίδευση και αξιολόγηση του μοντέλου πρόβλεψης. Τα δεδομένα των δημοφιλέστερων κρυπτονομισμάτων Proof-of-Work ξεκινούν από τα τέλη του 2014. Παρόλα αυτά, οι υπόλοιποι αλγόριθμοι ομοφωνίας αναπτύχθηκαν μεταγενέστερα, έχοντας ως αναφορά τους αλγορίθμους Proof-of-Work και σκοπό την βελτίωση των μειονεκτημάτων αυτών. Το σημαντικότερο Proof-of-Stake κρυπτονομίσμα που πρέπει να συμπεριληφθεί στην μελέτη, το Ethereum, δημιούργησε το genesis block του τον Ιούλιο του 2015, και τα πιο πρόσφατα χρηματιστηριακά δεδομένα αυτού (από το Yahoo Finance) ξεκινούν από τον Σεπτέμβριο του 2017. Έτσι η προαναφερθείσα ημερομηνία αποτελεί την παλαιότερη δυνατή χρονική στιγμή άντλησης για το μοντέλο. Παρόλα αυτά, μιας και είναι θεμιτό να συμπεριληφθούν και οι κατηγορίες των αλγορίθμων Proof-of-Authority και Federated Byzantine Agreement στην εργασία το διάστημα αυτό πρέπει να περιέχει και κάποια κρυπτονομίσματα αυτών των τεχνολογιών. Η παλαιότερη δυνατή ημερομηνία για κρυπτονομίσματα τύπου PoA είναι η 6^η Νοεμβρίου 2017, που είναι η ημερομηνία της πρώτης χρηματιστηριακής τιμής (Yahoo Finance) για το δημοφιλέστερο κρυπτονομίσμα PoA, το Binance Coin. Παρόλα αυτά προτιμήθηκε να συμπεριληφθεί και το δεύτερο πιο δημοφιλές PoA κρυπτονομίσμα, το VeChain, για πληρότητα των δεδομένων και έτσι τα δεδομένα ξεκινούν 10 μήνες μετά από την 3^η Αυγούστου 2018 έως την 8^η Σεπτεμβρίου 2023. Η συγκεκριμένη αρχική ημερομηνία δίνει την δυνατότητα να συμπεριληφθούν πολλά από τα πιο δημοφιλή κρυπτονομίσματα χωρίς να μειωθεί σημαντικά το χρονικό διάστημα των δεδομένων του μοντέλου και έτσι αυτή ορίστηκε ως η αρχή των δεδομένων.

Η επιλογή των κρυπτονομισμάτων έγινε με βάση την κεφαλαιοποίηση της αγοράς από το coinmarketcap.com εξαιρώντας όμως κρυπτονομίσματα που θα μείωναν σημαντικά το υποσύνολο τιμών μας. Έτσι για παράδειγμα, για τα κρυπτονομίσματα PoS δεν

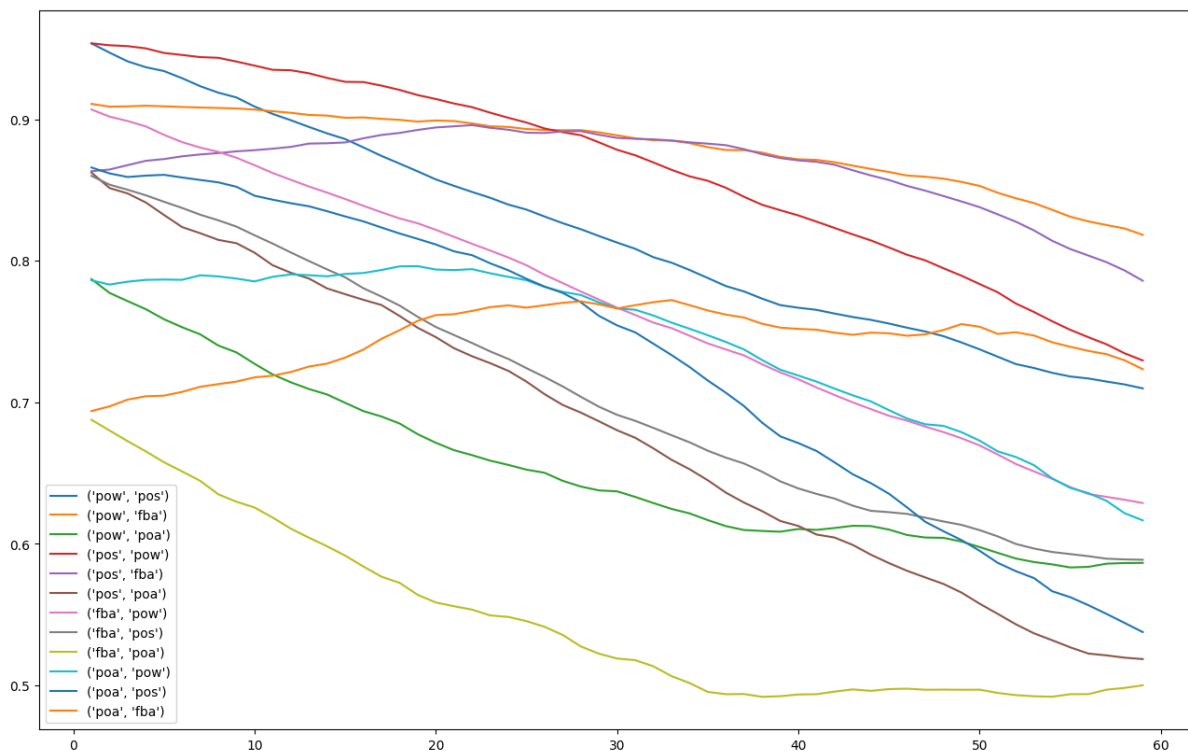
χρησιμοποιήθηκαν τα PolkaDot και Solana, επειδή τα διαθέσιμα δεδομένα τους ξεκινούσαν από το 2020 καταλήγοντας έτσι στο μεγαλύτερο δυνατό υποσύνολο τιμών, πράγμα αρκετά σημαντικό για την ύπαρξη επαρκών δεδομένων για το μοντέλο. Πιο συγκεκριμένα τα δεδομένα αποτελούν τις ημερήσιες τιμές κλεισίματος του χρηματιστηρίου για τα:

- **POW:** Bitcoin, Dash (Hybrid), Dogecoin, Litecoin, Monero
- **POS:** Ethereum, Cardano, Dash, Tron, Tezos
- **FBA:** Stellar, Ripple
- **POA:** Binance Coin, VeChain

Για τα δεδομένα επιλέχθηκε η ημερήσια τιμή κλεισίματος και όχι η ωριαία τιμή των κρυπτονομισμάτων, μετά από αναζήτηση του καλύτερου διαστήματος συσχέτισης σε σχέση με τον συντελεστή συσχέτισης Pearson.



Εικόνα 15. Συντελεστές συσχέτισης ωριαίων δεδομένων των μέσων δεικτών αλγορίθμων ομοφωνίας



Εικόνα 16. Συντελεστές συσχέτισης ημερήσιων δεδομένων των μέσων δεικτών αλγορίθμων ομοφωνίας

Οι παραπάνω γραφικές παραστάσεις παρουσιάζουν τους συντελεστές συσχέτισης των μέσων δεικτών αλγορίθμων ομοφωνίας (ο υπολογισμός τους αναφέρεται παρακάτω) για δεδομένα στο ίδιο χρονικό διάστημα (9 Σεπτεμβρίου 2021 – 8 Σεπτεμβρίου 2023) τόσο σε ωριαία όσο και σε ημερήσια δεδομένα. Ο άξονας y της γραφικής παράστασης αναπαριστά τον συντελεστή συσχέτισης, και ο άξονας x είναι το διάστημα ολίσθησης (lag) του ενός δείκτη σε σχέση με τον άλλο, σε ώρες (στο πρώτο διάγραμμα), και σε μέρες (στο δεύτερο). Έτσι για παράδειγμα στα δύο διαγράμματα, η κόκκινη γραμμή δείχνει τον συντελεστή συσχέτισης του μέσου δείκτη PoS με τον μέσο δείκτη PoW καθώς ο δεύτερος μετατοπίζεται κατά x ώρες/μέρες αναλόγως το διάγραμμα, προς το παρελθόν. Πρακτικά, προσπαθούμε να χρησιμοποιήσουμε παρελθοντικές τιμές του δείκτη PoW για να προβλέψουμε μελλοντικές τιμές του δείκτη PoS και είναι λογικό πως όσο καλύτερη είναι η γραμμική συσχέτιση τόσο καλύτερη θα είναι και η προβλεπτική ικανότητα με βάση τον επιλεγμένο δείκτη.

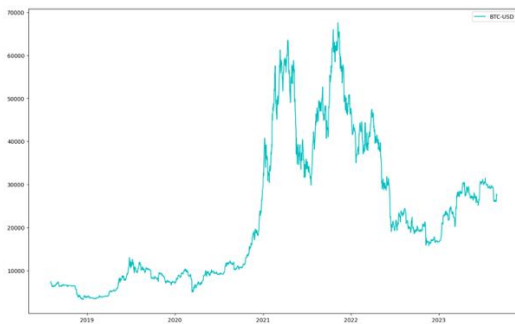
Για τα ημερήσια δεδομένα οι τιμές του συντελεστή συσχέτισης ξεκινούν από το 0.5, στην χειρότερη, σε μεγαλύτερα lags, και αυξάνονται έως τιμές κοντά στο 1, την μέγιστη τιμή του συντελεστή. Από την άλλη, τα ωριαία δεδομένα έχουν ως μέγιστη τιμή το 0.4. Είναι λοιπόν εύκολα κατανοητό ότι τα ημερήσια δεδομένα έχουν καλύτερη γραμμική συσχέτιση σε σχέση με τα ωριαία, και γι' αυτό επιλέχθηκαν τα ημερήσια δεδομένα για την εκπαίδευση και αξιολόγηση του μοντέλου.

Παράλληλα λόγω περιορισμών τόσο του Yahoo Finance όσο και του coinmarketcap, όπου στο πρώτο το μέγιστο διάστημα άντλησης ωριαίων δεδομένων είναι τα δύο τελευταία χρόνια, ενώ στο δεύτερο δεν υπήρχε δωρεάν διαθέσιμο πλάνο για την άντληση ιστορικών δεδομένων, προτιμήθηκαν πάλι τα ημερήσια δεδομένα, στα οποία δεν υπήρχε τέτοιος περιορισμός (στο Yahoo Finance). Ειδικότερα εξαιτίας της πανδημίας του κορονοϊού, τα δεδομένα για τα κρυπτονομίσματα μετά το 2020 θα

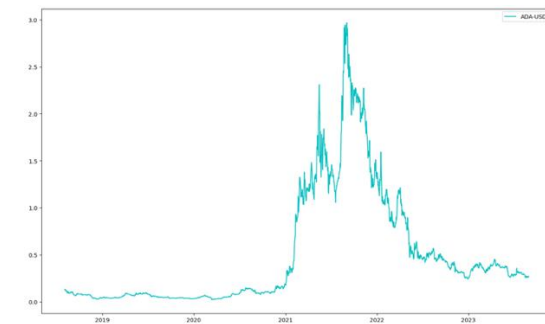
παρουσίαζαν μια γενική πτωτική τάση που αντικατοπτρίζεται γενικά σε όλο το χρηματιστήριο καθώς και μεγαλύτερη αστάθεια σε σχέση με πριν το 2020, οπότε και για αυτό τον λόγο επιλέχθηκε να συμπεριληφθούν και παλαιότερα δεδομένα ξεκινώντας από το 2018. Τα τελικά δεδομένα συλλέχθηκαν από το Yahoo Finance, μέσω της βιβλιοθήκης “yfinance” της Python.

Στην τελική μορφή του μοντέλου, και με βάση την αντίστοιχη βιβλιογραφία, αποφασίστηκε να μελετηθεί και η σχέση με κάποιους βασικούς χρηματιστηριακούς δείκτες που είναι το πετρέλαιο, ο χρυσός και ο χρηματιστηριακός δείκτης S&P500, για να μελετήσουμε την επιρροή τους, και έτσι την άμεση σχέση τους με τα κρυπτονομίσματα, στην τελική πρόβλεψη. Τα δεδομένα για τα τρία αυτά μεγέθη όπως και για τα κρυπτονομίσματα αντλούνται με τον ίδιο τρόπο από το Yahoo Finance από την ίδια ημερομηνία (3^η Αυγούστου 2018) και αναφέρονται και αυτά σε ημερήσιες τιμές κλεισίματος.

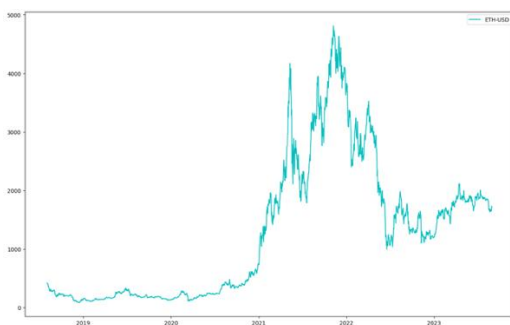
Όλα τα παραπάνω δεδομένα χρησιμοποιήθηκαν τόσο για την εκπαίδευση όσο και για την αξιολόγηση του πειραματικού μοντέλου όπως θα αναλυθεί παρακάτω. Κάποιες ενδεικτικές τιμές από τα δημοφιλέστερα κρυπτονομίσματα παρουσιάζονται ως εξής.



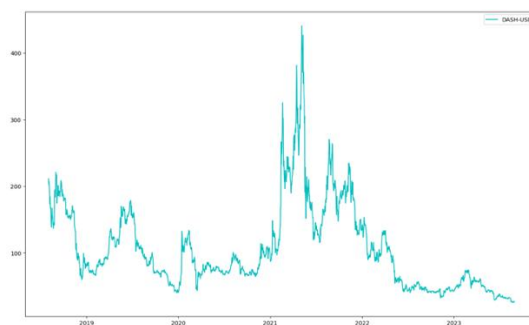
Εικόνα 17. Ιστορικές τιμές του Bitcoin



Εικόνα 18. Ιστορικές τιμές του Cardano



Εικόνα 19. Ιστορικές τιμές του Ethereum

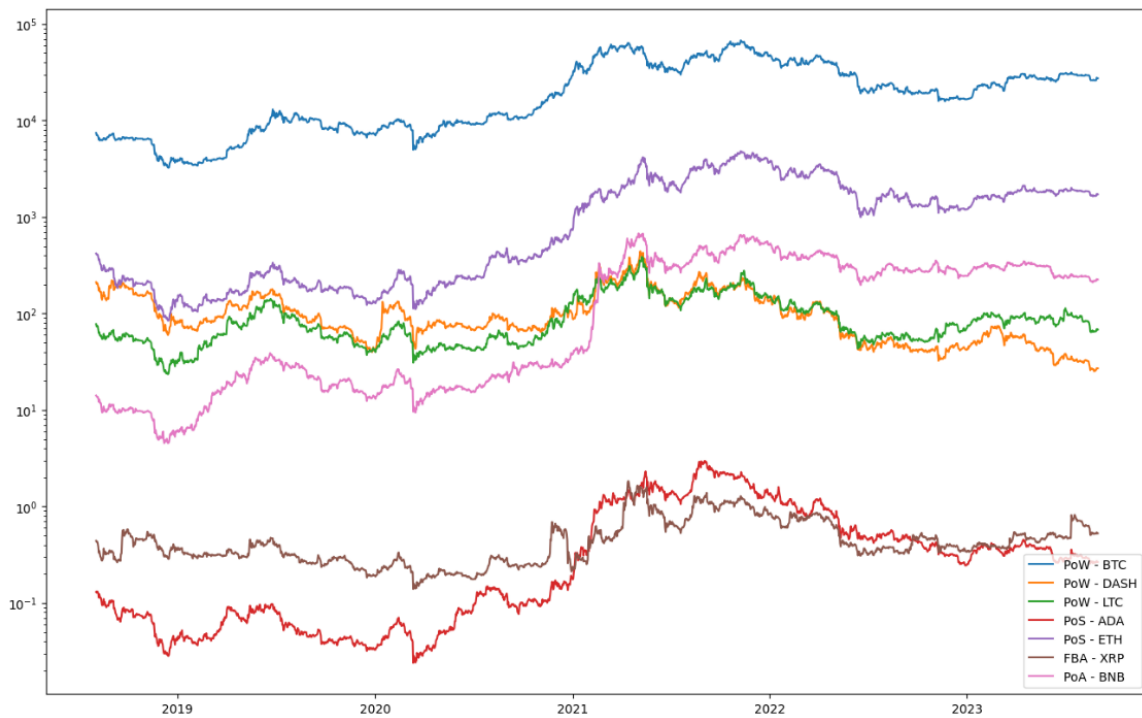


Εικόνα 20. Ιστορικές τιμές του Dash

5.2 Σύγκριση των αρχικών δεδομένων

Από τις παραπάνω γραφικές είναι εύκολα κατανοητό, ότι σε παρόμοια χρονικά διαστήματα τα κρυπτονομίσματα ακολουθούν παρόμοιες τάσεις αύξησης και μείωσης. Αυτό το συμπέρασμα επαληθεύεται και από την γραφική παράσταση των συντελεστών συσχέτισης Pearson, όπου και παρουσιάζονται τιμές πολύ κοντά στην μονάδα (πράγμα που φανερώνει ισχυρή γραμμική συσχέτιση) και οι περισσότεροι δείκτες εμφανίζουν

τις μέγιστες τιμές συντελεστών Pearson όσο τείνει το διαστήματα ολίσθησης (lags) στο μηδέν. Η πορεία μεμονωμένων κρυπτονομισμάτων φαίνεται στην παρακάτω γραφική.



Εικόνα 21. Τιμές κρυπτονομισμάτων ανά τον αλγόριθμο ομοφωνίας

Σε αυτή την γραμμική συσχέτιση θα βασιστεί η ανάλυση για την εξαγωγή συμπερασμάτων και πρόβλεψη τιμής που θέλουμε να επιτευχθεί. Από την άλλη η σχέση του Bitcoin με τους τρεις δείκτες που θα χρησιμοποιήσουμε παρακάτω για την επέκταση του μοντέλου είναι αρκετά διαφορετική.



Εικόνα 22. Σύγκριση του Bitcoin με άλλους χρηματιστηριακούς δείκτες

Σε σύγκριση με τα υπόλοιπα κρυπτονομίσματα, οι χρηματιστηριακοί δείκτες έχουν μικρότερη συσχέτιση με το Bitcoin. Παρατηρούμε κάποια μικρές αυξομειώσεις τόσο στους τρεις δείκτες όσο και στο Bitcoin που συμβαίνουν ταυτόχρονα, που πιθανώς οφείλονται σε στιγμιαία πτώση/αύξηση της αγοράς. Παράλληλα παρατηρείται και μία μικρή αυξητική τάση που ακολουθούν τόσο οι δείκτες όσο και το Bitcoin, αλλά όχι στον βαθμό συσχέτισης που παρατηρήθηκε μεταξύ των ίδιων των κρυπτονομισμάτων. Στην γραφική φαίνεται εύκολα και η πιο ασταθής φύση του Bitcoin, με έντονες αυξομειώσεις σε σχέση με τους χρηματιστηριακούς δείκτες που έχουν μία σταθερή αυξητική τάση και μικρές πτώσεις της τιμής τους.

5.3 Επεξεργασία δεδομένων και δημιουργία γενικών δεικτών αλγορίθμων ομοφωνίας

Ένας από τους στόχους της διπλωματικής εργασίας, είναι η μελέτη της σχέσης των κρυπτονομισμάτων μέσω κάποιων γενικών χαρακτηριστικών τους, και η ανάλυση των παραγόντων αυτών. Το σημαντικότερο χαρακτηριστικό που θα μελετηθεί είναι οι αλγόριθμοι ομοφωνίας και πως αυτοί σχετίζονται μεταξύ τους.

Για την μελέτη αυτή δημιουργούμε κάποιους γενικούς δείκτες για κάθε αλγόριθμο ομοφωνίας. Ο κάθε δείκτης για τον συγκεκριμένο αλγόριθμο ομοφωνίας, αποτελείται από τον σταθμισμένο μέσο όρο όλων των κρυπτονομισμάτων της ίδιας τεχνολογίας χρησιμοποιώντας ως βάρος τον χρηματιστηριακό τους όγκο. Η μαθηματική σχέση είναι η εξής.

$$WMA_{consensus} = \sum_i^n \sum_t^T \frac{crypto_price_i(t) * crypto_volume_i(t)}{\sum_j^n crypto_volume_j(t)}$$

Όπου *crypto_price* είναι η τιμή του κρυπτονομίσματος την χρονική στιγμή *t*, και *crypto_volume* ο όγκος του συγκεκριμένου κρυπτονομίσματος την ίδια χρονική στιγμή. Έτσι για παράδειγμα για τον γενικό δείκτη του αλγορίθμου Proof-of-Work θα έχουμε.

$$PoW = \sum \frac{BTC(t) * V_{BTC}(t)}{V_{total}(t)} + \sum \frac{DASH(t) * V_{DASH}(t)}{V_{total}(t)} + \sum \frac{LTC(t) * V_{LTC}(t)}{V_{total}(t)} + \sum \frac{DOGE(t) * V_{DOGE}(t)}{V_{total}(t)} + \sum \frac{XMR(t) * V_{XMR}(t)}{V_{total}(t)}$$

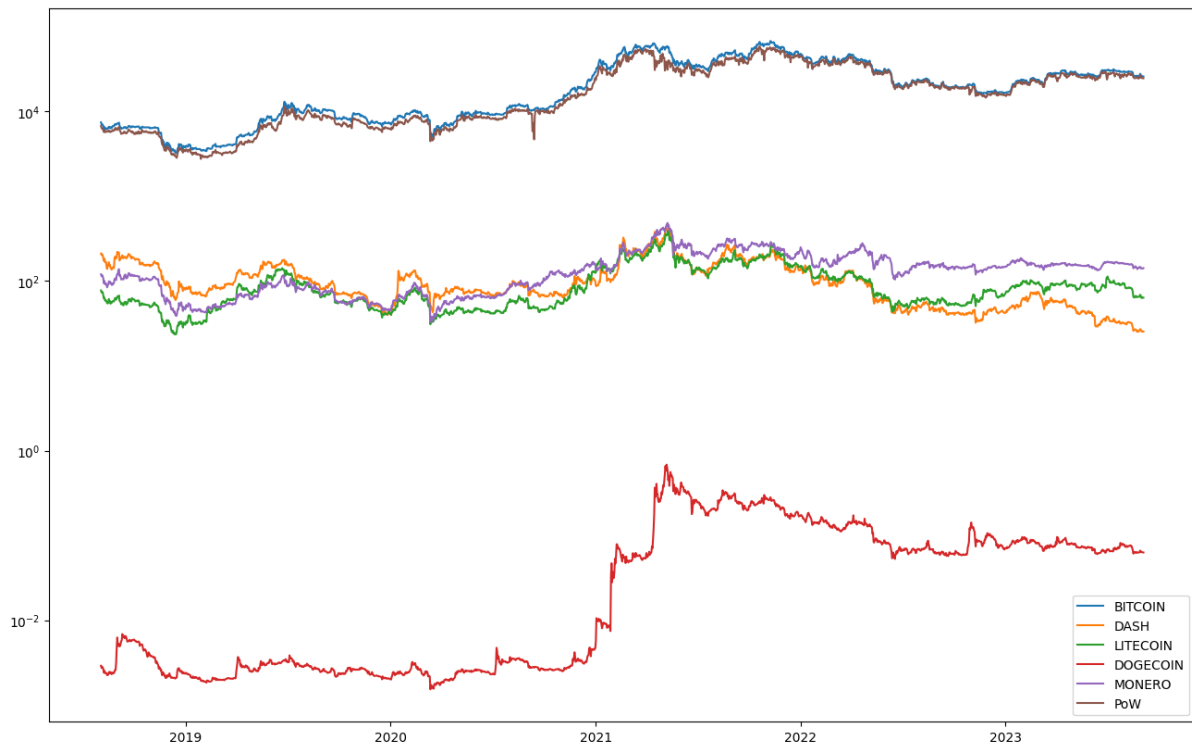
Όπου V_{total} είναι ο συνολικός όγκος όλων των κρυπτονομισμάτων την χρονική στιγμή *t*.

$$V_{total}(t) = V_{BTC}(t) + V_{DASH}(t) + V_{LTC}(t) + V_{DOGE}(t) + V_{XMR}(t)$$

Οι συγκεκριμένοι δείκτες είναι πιο αντιπροσωπευτικοί σε σχέση με τον απλό μέσο όρο για τα κρυπτονομίσματα, γιατί τα κρυπτονομίσματα με χαμηλότερη τιμή συνεισφέρουν περισσότερο στον σταθμισμένο μέσο όρο. Για παράδειγμα, για τον γενικό δείκτη PoW λόγω της τεράστιας τιμής του Bitcoin, σε σχέση με όλα τα υπόλοιπα νομίσματα, ο τελικός δείκτης θα ήταν πρακτικά η τιμή του ίδιου του Bitcoin, αλλά τώρα λόγω του βάρους του όγκου, τα υπόλοιπα κρυπτονομίσματα συμμετέχουν σε μεγαλύτερο βαθμό.

Βέβαια παρόλο που ο σταθμισμένος μέσος όρος παρέχει μεγαλύτερη συνεισφορά στα κρυπτονομίσματα με μικρότερη τιμή, σε περιπτώσεις όπως αυτές του Bitcoin, όπου η

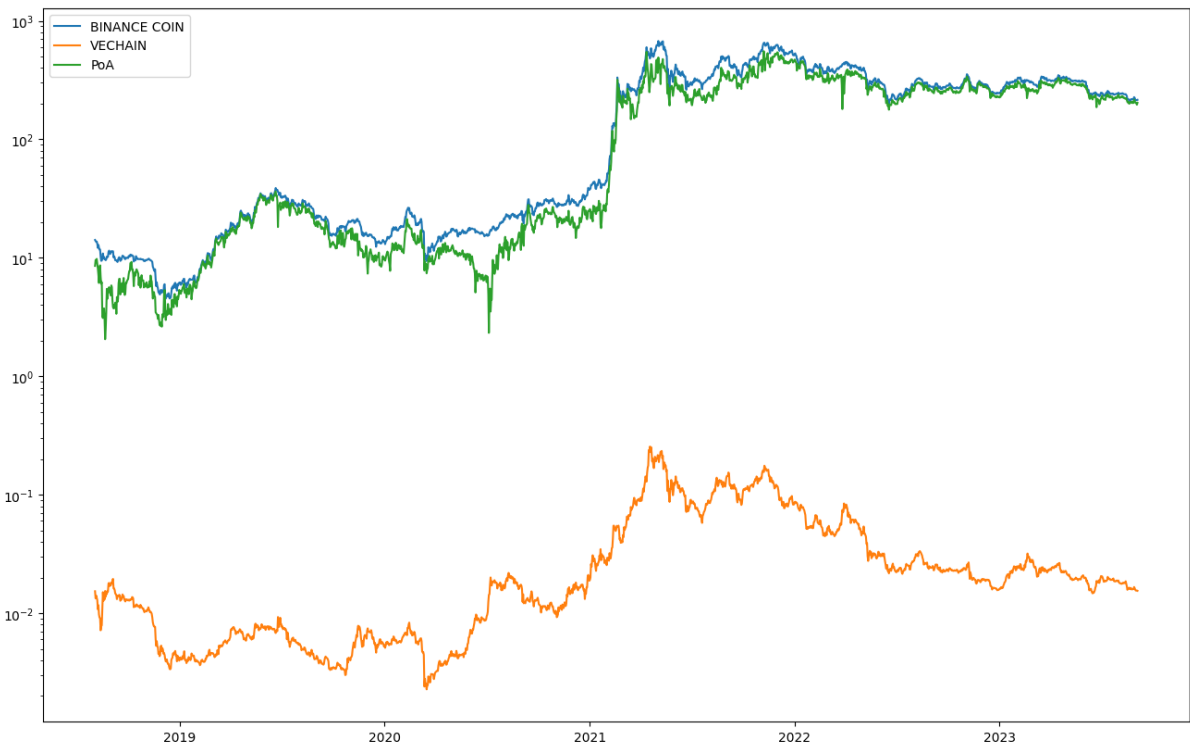
τιμή είναι σημαντικά μεγαλύτερη, ο δείκτης πλησιάζει πολύ κοντά στο κρυπτονόμισμα με την μεγαλύτερη τιμή. Αυτό όμως δεν δημιουργεί πρόβλημα, γιατί θεωρητικά τουλάχιστον, τα κρυπτονομίσματα με την μεγαλύτερη τιμή, επηρεάζουν πιο σημαντικά την τιμή των υπολοίπων, εξαιτίας της δημοτικότητας τους.



Εικόνα 23. Ο γενικός δείκτης PoW σε σύγκριση με τα επιμέρους κρυπτονομίσματα



Εικόνα 24. Ο γενικός δείκτης PoS σε σύγκριση με τα επιμέρους κρυπτονομίσματα



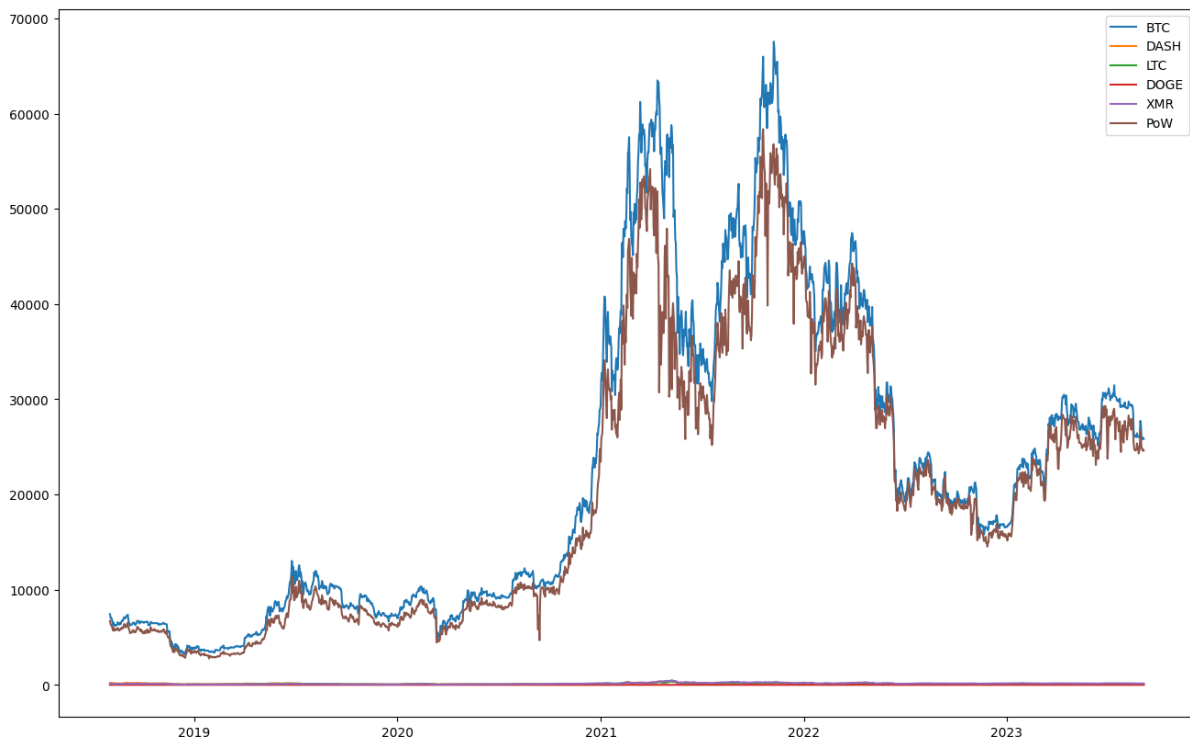
Εικόνα 25. Ο γενικός δείκτης PoA σε σύγκριση με τα επιμέρους κρυπτονομίσματα



Εικόνα 26. Ο γενικός δείκτης FBA σε σύγκριση με τα επιμέρους κρυπτονομίσματα

Τελικά, όπως φαίνεται από τις παραπάνω γραφικές, όλοι οι γενικοί δείκτες πλησιάζουν το κρυπτονόμισμα με την μεγαλύτερη τιμή, πράγμα αναπόφευκτο, γιατί σε κάθε κατηγορία κρυπτονομισμάτων ένα τουλάχιστον κρυπτονόμισμα έχει τιμές σημαντικά μεγαλύτερες από τα άλλα. Παρόλα αυτά εξαιτίας του τρόπου που ορίσαμε τον δείκτη, ακόμα και τα μικρά κρυπτονομίσματα έχουν σημαντική διαφορά στην τελική τιμή του

δείκτη, πράγμα που μπορεί να φανεί για παράδειγμα, αν αφαιρέσουμε την λογαριθμική κλίμακα στην γραφική παράσταση του δείκτη του PoW.



Εικόνα 27. Ο γενικός δείκτης PoW σε σχέση με το Bitcoin

Ο στόχος μας κατά την δημιουργία των δεικτών είναι η αναπαράσταση μίας μέσης συμπεριφοράς της τεχνολογίας, και όπως βλέπουμε από τα παραπάνω διάγραμμα αυτό έχει επιτευχθεί σε αρκετά καλό βαθμό. Ένα τέτοιο παράδειγμα είναι η απότομη πτώση (dip) σε σχέση με το κύριο κρυπτονόμισμα σε όλους τους δείκτες, πράγμα που οφείλεται σε σημαντική πτώση των επιμέρους κρυπτονομισμάτων. Σε μία τέτοια περίπτωση θα περιμένουμε το μοντέλο να προβλέψει γενική πτώση στην τελική πρόβλεψη γιατί η γενική αγορά έχει επηρεαστεί αρνητικά παρόλο που το σημαντικό κρυπτονόμισμα κάθε φορά – εξαιτίας της μεγάλης τιμής του – παραμένει πιο σταθερό. Μία επιπλέον παρατήρηση είναι ότι οι μέσοι δείκτες έχουν πάντα μικρότερη τιμή από το σημαντικότερο κρυπτονόμισμα, πράγμα λογικό γιατί τα επιπλέον κρυπτονομίσματα μειώνουν την τιμή του δείκτη. Παρόλα αυτά, απότομες αυξήσεις στον γενικό δείκτη που τείνουν να φτάσουν την τιμή του κύριου κρυπτονομίσματος δείχνουν μία σημαντική αύξηση της τιμής των επιμέρους κρυπτονομισμάτων. Γενικά οι δείκτες που δημιουργήθηκαν προσπαθούν να παράξουν μεγαλύτερη σταθερότητα λαμβάνοντας υπόψιν πολλά και διάφορα κρυπτονομίσματα για την λήψη αποφάσεων.

Παράλληλα, όπως αναφέρθηκε και παραπάνω, οι δείκτες χρησιμοποιούνται και για την εύρεση του καλύτερου διαστήματος ολίσθησης (lag) για την πρόβλεψη της αντίστοιχης τεχνολογίας κάθε φορά. Παρακάτω δίνεται ο πίνακας του διαστήματος των πρώτων 10 ημερών ολίσθησης του κάθε δείκτη.

	PoW			PoS			FBA			PoA		
	PoS	FBA	PoA	PoW	FBA	PoA	PoW	PoS	PoA	PoW	PoS	FBA
1	0.912	0.785	0.851	0.915	0.795	0.942	0.792	0.792	0.736	0.855	0.943	0.736
2	0.910	0.781	0.847	0.915	0.796	0.940	0.793	0.788	0.733	0.856	0.942	0.737
3	0.907	0.777	0.842	0.915	0.797	0.938	0.795	0.786	0.729	0.858	0.941	0.737
4	0.905	0.775	0.838	0.915	0.797	0.937	0.796	0.784	0.725	0.859	0.940	0.736
5	0.903	0.772	0.834	0.915	0.797	0.935	0.797	0.780	0.722	0.860	0.940	0.736
6	0.900	0.770	0.832	0.915	0.798	0.933	0.797	0.777	0.718	0.862	0.939	0.737
7	0.897	0.768	0.829	0.916	0.799	0.932	0.797	0.773	0.714	0.864	0.938	0.737
8	0.895	0.766	0.825	0.917	0.800	0.932	0.798	0.769	0.709	0.866	0.937	0.736
9	0.892	0.766	0.823	0.917	0.801	0.931	0.798	0.765	0.705	0.867	0.936	0.735
10	0.890	0.764	0.820	0.918	0.802	0.930	0.798	0.761	0.700	0.869	0.935	0.734

Πίνακας 3. Παράδειγμα συντελεστών συσχέτισης των γενικών δεικτών

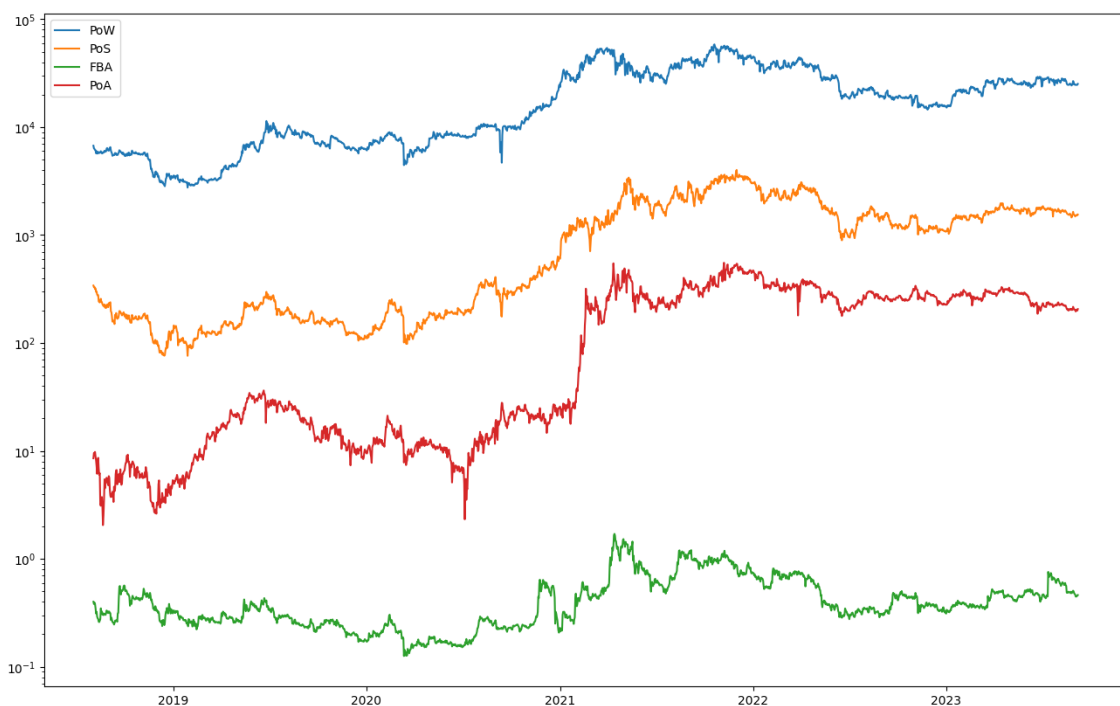
Από τον παραπάνω πίνακα θα αντλήσουμε κατά την δημιουργία του μοντέλου τα καλύτερα διαστήματα ολίσθησης για την τροφοδοσία του μοντέλου με τους δείκτες που παρουσιάζουν την καλύτερη γραμμική συσχέτιση. Οι συντελεστές αυτοί παρουσιάστηκαν και παραπάνω σχηματικά στην συλλογή των δεδομένων.

Κεφάλαιο 6. Μελέτη του μοντέλου και ανάλυση αποτελεσμάτων

Στόχος της διαδικασίας της μελέτης είναι η άντληση συμπερασμάτων για την επιρροή των υποκείμενων τεχνολογιών των κρυπτονομισμάτων στην χρηματιστηριακή τιμή τους ή χρήση των συμπερασμάτων αυτών για πρόβλεψη των τιμών και σε επόμενο στάδιο η ανάλυση της σχέσης των διάφορων αυτών τεχνολογιών μεταξύ τους. Πιο συγκεκριμένα, ο κυριότερος τεχνολογικός παράγοντας που διαφοροποιεί τα κρυπτονομίσματα είναι ο αλγόριθμος ομοφωνίας τους, όπως αναφέρθηκε και στα προηγούμενα κεφάλαια. Ο αλγόριθμος αυτός είναι ο κύριος πυλώνας διαχώρισης των δεδομένων που θα χρησιμοποιηθούν. Παρόλα αυτά επειδή το κάθε κρυπτονομίσμα έχει διαφοροποιήσεις στον τρόπο ανάθεσης κόμβων, στον αλγόριθμο κρυπτογράφησης, στην επιβράβευση των κόμβων, κλπ., είναι αδύνατη η εξουχιστική σύγκριση όλων αυτών των παραγόντων και έτσι η διαχώριση των δεδομένων γίνεται με βάση κάποιες βασικές δομικές αρχές. Η διαχώριση αυτή πρακτικά είναι οι βασικές κατηγορίες των αλγορίθμων ομοφωνίας που ήδη αναφέρθηκαν στο δεύτερο κεφάλαιο. Έτσι δημιουργήθηκαν οι τέσσερις γενικοί δείκτες PoW (Proof-of-Work), PoS (Proof-of-Stake), PoA (Proof-of-Authority), FBA (Federated Byzantine Agreement) που παρουσιάστηκαν στο προηγούμενο κεφάλαιο, που θα αποτελέσουν τα δεδομένα του μοντέλου.

6.1 Διαχωρισμός δεδομένων

Μετά από την επεξεργασία και την δημιουργία των τεσσάρων γενικών δεικτών αλγορίθμου ομοφωνίας, όπως αναφέρθηκε στο πέμπτο κεφάλαιο, καταλήγουμε στα δεδομένα που θα αποτελέσουν την είσοδο του μοντέλου. Αυτά είναι τέσσερις χρονοσειρές των γενικών δεικτών αλγορίθμου ομοφωνίας με αρχή την 3^η Αυγούστου 2018 και τέλος την 8^η Σεπτεμβρίου 2023, σε αναλογία με την επιλογή των ίδιων χρονικών διαστημάτων στα μεμονωμένα κρυπτονομίσματα.



Εικόνα 28. Τέσσερις γενικοί δείκτες αλγορίθμου ομοφωνίας

Οι χρονοσειρές που προκύπτουν είναι συνεχείς, χωρίς να περιέχουν μηδενικές τιμές, έτσι ώστε να αποφευχθούν τυχόν σφάλματα στους δείκτες ακρίβειας. Οι χρονοσειρές αυτές θα διαχωριστούν σε δύο κομμάτια όπου το ένα θα χρησιμοποιηθεί για την εκπαίδευση του μοντέλου, και το άλλο για την αξιολόγηση αυτού. Για τον διαχωρισμό αυτό ακολουθήθηκε ο άτυπος κανόνας του 80/20, όπου τεμαχίζεται το σύνολο δεδομένων με τέτοιο τρόπο ώστε περίπου το 20% αυτού να αποτελεί δεδομένα αξιολόγησης. Στην δική μας περίπτωση οι τελευταίοι 12 μήνες των δεδομένων χρησιμοποιήθηκαν ως το σύνολο αξιολόγησης από την 1^η Οκτωβρίου 2022, που είναι περίπου το 20% του συνολικού πλήθους δεδομένων.

Το μοντέλο που θα δημιουργήσουμε, έχει στόχο να βρει την σχέση μεταξύ των διαφορετικών τεχνολογιών των κρυπτονομισμάτων. Έτσι για να γίνει αυτή η σύγκριση, σαν είσοδο κάθε φορά δίνεται το σύνολο των δεδομένων εκπαίδευσης τριών από τις τέσσερις χρονοσειρές των δεικτών. Στόχος είναι να προβλεφθούν οι τιμές της τέταρτης χρονοσειράς, η οποία δεν επιλέχθηκε κάθε φορά. Από την στιγμή που θέλουμε να προβλέψουμε τιμές, πρέπει να χρησιμοποιήσουμε παρελθοντικές παρατηρήσεις των δεδομένων στα οποία θα βασιστούν οι προβλέψεις. Έτσι το πρώτο βήμα της εκπαίδευσης, είναι να βρεθεί η καλύτερη γραμμική συσχέτιση της κάθε χρονοσειράς εισόδου με την χρονοσειρά που θέλουμε να προβλέψουμε. Για να πάρουμε τις παρελθοντικές τιμές πρέπει να μετατοπίσουμε την κάθε χρονοσειρά κατά x βήματα προς το παρελθόν, αφαιρώντας τις x πιο πρόσφατες τιμές. Στο τέλος καταλήγουμε με ένα υποσύνολο της αρχικής χρονοσειράς, όπου το πρώτο δεδομένο χρησιμοποιείται για να προβλέψει το δεδομένα x μέρες μετά της χρονοσειράς που θέλουμε να προβλέψουμε. Είναι εύκολα κατανοητό σε αυτό το σημείο, ότι λόγω του πλήθους των δεδομένων που αφαιρούνται, δεν μπορούν να προβλεφθούν οι τελευταίες x τιμές της χρονοσειράς μιας και δεν έχουμε μελλοντικά δεδομένα των χρονοσειρών εισόδου, αλλά αυτό δεν αποτελεί πρόβλημα. Το μοντέλο παρόλα αυτά χρησιμοποιεί το σύνολο των τριών χρονοσειρών σαν είσοδο με διαφορετικές ολισθήσεις x , y , z . Το τελικό μοντέλο δεν μπορεί να προβλέψει τις τελευταίες x ή y ή z τιμές, αναλόγως του ποια από τις 3 μεταβλητές παίρνει την μεγαλύτερη τιμή.

Η επιλογή των τιμών x , y και z γίνεται με βάση την καλύτερη γραμμική συσχέτιση που έχουν οι χρονοσειρές. Ο υπολογισμός αυτός γίνεται μετατοπίζοντας την κάθε χρονοσειρά κατά μία μέρα την φορά, και μετά υπολογίζεται ο συντελεστής συσχέτισης Pearson των δύο χρονοσειρών. Η διαδικασία αυτή επαναλαμβάνεται για ένα διάστημα 60 ημερών και για τις τρεις χρονοσειρές. Ένα παράδειγμα των αποτελεσμάτων που προκύπτουν είναι η γραφική παράσταση των συντελεστών συσχέτισης της εικόνας 16 και ο πίνακας 3 του πέμπτου κεφαλαίου, όπου για τον πίνακα, η πρώτη χρονοσειρά είναι αυτή που θα προβλεφθεί, η δεύτερη είναι η μετατοπισμένη κατά x μέρες, και η κάθε γραμμή είναι η μετατόπιση σε μέρες κάθε φορά. Ο τελικός πίνακας που υπολογίζεται έχει 60 γραμμές, και από αυτές παίρνουμε την γραμμή, και άρα την μετατόπιση, με τον μεγαλύτερο συντελεστή συσχέτισης. Ο τρόπος που χρησιμοποιούνται οι τελικές χρονοσειρές θα παρουσιαστεί στο επόμενο κεφάλαιο. Τα τελικά βέλτιστα διαστήματα ολίσθησης είναι:

Βέλτιστα διαστήματα ολίσθησης				
	PoW - shifted	PoS - shifted	PoA - shifted	FBA - shifted
PoW		1	1	1
PoS	35		1	20
PoA	34	1		7
FBA	24	1	1	

Πίνακας 4. Βέλτιστα διαστήματα ολίσθησης

6.2 Επιλογή του βέλτιστου μοντέλου

Το μοντέλο που θα χρησιμοποιήσουμε είναι αυτό της γραμμικής παλινδρόμησης, τροφοδοτώντας το με τις χρονοσειρές των αλγορίθμων ομοφωνίας οι οποίες προέκυψαν με τον τρόπο που αναφέρθηκε παραπάνω. Το μοντέλο της γραμμικής παλινδρόμησης επιλέχθηκε, τόσο για την ευκολία χρήσης του όσο και για την ευκολία σύγκρισης των αποτελεσμάτων. Στο αρχικό στάδιο θα δοθούν τα δεδομένα εισόδου στην πιο απλή μορφή τους, και σε βήματα θα προστεθούν διαφοροποιήσεις των αρχικών δεδομένων. Σε κάθε βήμα θα γίνει αξιολόγηση των αποτελεσμάτων του μοντέλου με βάση κάποιους δείκτες ακρίβειας, για να βρεθεί το καλύτερο σύνολο δεδομένων εισόδου και να βρεθεί το βέλτιστο τελικό μοντέλο. Σε αυτή την διαδικασία, αν χρησιμοποιούνταν κάποιο πιο πολύπλοκο μοντέλο θα ήταν δυσκολότερη η συσχέτιση των αποτελεσμάτων σε κάθε βήμα. Παρόλα αυτά σαν επέκταση, η χρήση του τελικού συνόλου δεδομένων με κάποιο καλύτερο και πολυπλοκότερο μοντέλο πιθανώς θα έδινε καλύτερα αποτελέσματα, αλλά ο κύριος στόχος της παρούσας διπλωματικής εργασίας είναι η εύρεση των σχέσεων των τεχνολογιών και σε δεύτερο βαθμό η ποιότητα των τελικών προβλέψεων. Τέλος, καθώς στα δεδομένα παρατηρήθηκε μεγάλη γραμμική συσχέτιση, θεωρήθηκε ότι από τα πιο απλά μοντέλα που αναφέρθηκαν στο τέταρτο κεφάλαιο, θα ήταν από τις καλύτερες επιλογές ώστε να παραχθούν επαρκώς καλές προβλέψεις με βάση τους στόχους στα πλαίσια της εργασίας.

6.2.1 Μοντέλο με χρήση βέλτιστων ολισθήσεων

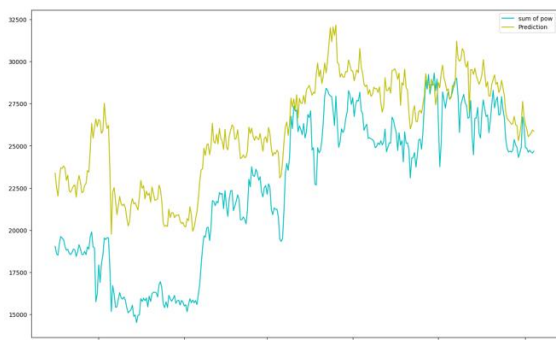
Το πρώτο μοντέλο που παρουσιάζεται είναι αυτό που χρησιμοποιεί σαν δεδομένα εισόδου τις τρεις χρονοσειρές των αλγορίθμων ομοφωνίας, μετατοπισμένες τόσες μέρες ώστε να προκύψει η καλύτερη γραμμική συσχέτιση, όπως αναφέρθηκε παραπάνω.

Έτσι για παράδειγμα, για την πρόβλεψη της χρονοσειράς PoS, χρησιμοποιούνται οι χρονοσειρές PoW, PoA, FBA μετατοπισμένες κατά 35, 1 και 20 ημέρες αντίστοιχα με βάση τον πίνακα 4. Τα αποτελέσματα πρόβλεψης είναι τα εξής.

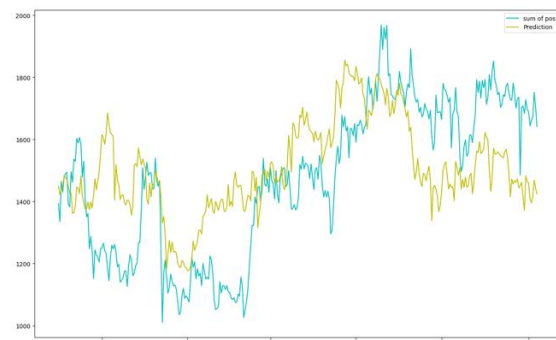
Πρωτόκολλο	RMSE	MAPE
PoW	3924.86	0.174424
PoS	199.516	0.122581
FBA	0.104668	0.225101
PoA	89.9932	0.295128

Παρόλης της απλότητας του πρώτου μοντέλου, μπορούν να εξαχθούν αρκετές πληροφορίες για την σχέση των τεχνολογιών. Παρατηρώντας τους δείκτες MAPE, βλέπουμε ότι οι διαφορετικές τεχνολογίες έχουν αρκετή απόκλιση μεταξύ τους, όσον αφορά τα σφάλματα προβλέψεων. Τα καλύτερα αποτελέσματα έχουν τα κρυπτονομίσματα που χρησιμοποιούν την τεχνολογία Proof-of-Stake, τα οποία παρουσιάζουν λιγότερο από το μισό ποσοστό σφάλματος των κρυπτονομισμάτων με τον χειρότερο δείκτη MAPE. Αυτή η διαφορά είναι μεγάλη και αρχικά δείχνει ότι τα κρυπτονομίσματα τύπου Proof-of-Stake παρουσιάζουν μεγάλη ομοιότητα και με τις τρεις άλλες κατηγορίες έχοντας αρκετά μικρό δείκτη σφάλματος ήδη από το απλούστερο μοντέλο. Παράλληλα τα αποτελέσματα δείχνουν ότι όντως υπάρχει διαφορά συμπεριφοράς με βάση τις διαφορετικές τεχνολογίες. Αν οι τεχνολογίες δεν έπαιζαν ρόλο στην εξέλιξη της τιμής τα αποτελέσματα θα περιμέναμε να είχαν μικρές αποκλίσεις σφαλμάτων μεταξύ τους, πράγμα που δεν συμβαίνει. Παρόλα αυτά οι δείκτες δεν δίνουν την ολοκληρωμένη εικόνα της σχέσης του μοντέλου με τις προβλέψεις, και θα χρησιμοποιηθούν κυρίως για την σύγκριση των διαφορετικών σταδίων των μοντέλων.

Για την αξιολόγηση των προβλέψεων θα μελετηθούν και οι γραφικές παραστάσεις των χρονοσειρών αλγορίθμων ομοφωνίας και των προβλέψεων, που δίνουν περισσότερα δεδομένα για την γραμμική συσχέτιση των προβλέψεων με των χρονοσειρών.



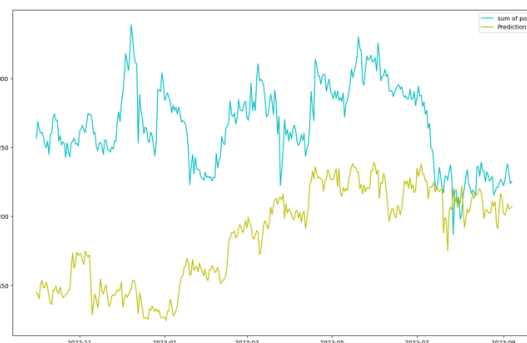
Εικόνα 29. Πρόβλεψη του δείκτη PoW με την χρήση του μοντέλου βέλτιστης ολίσθησης



Εικόνα 30. Πρόβλεψη του δείκτη PoS με την χρήση του μοντέλου βέλτιστης ολίσθησης



Εικόνα 31. Πρόβλεψη του δείκτη FBA με την χρήση του μοντέλου βέλτιστης ολίσθησης



Εικόνα 32. Πρόβλεψη του δείκτη PoA με την χρήση του μοντέλου βέλτιστης ολίσθησης

Από τις γραφικές παραστάσεις των προβλέψεων, παρατηρούμε ότι παρόλο που τον μικρότερο δείκτη MAPE τον έχει ο δείκτης PoS, την καλύτερη γραμμική συσχέτιση πρόβλεψης με πραγματική τιμή την παρουσιάζει ο δείκτης PoW. Σε αυτόν, οι προβλέψεις ακολουθούν πολύ κοντά τις απότομες διακυμάνσεις του δείκτη, απλά το γενικό επίπεδο των προβλέψεων απέχει περισσότερο από τον μέσο όρο των παρατηρήσεων σε σχέση με τον δείκτη PoS. Επίσης παρατηρούμε ότι παράλληλα με τον δείκτη MAPE, οι PoW και PoS που έχουν σχετικά μικρότερες τιμές σφάλματος έχουν και σχετικά καλή γραμμική συσχέτιση όσον αφορά τις προβλέψεις. Από την άλλη οι δείκτες PoA, FBA δεν έχουν υψηλή γραμμική συσχέτιση γραφικά. Η πρόβλεψη του FBA ακολουθεί μια ανοδική τάση όπως και οι παρατηρήσεις, αλλά έχει μεγάλες διακυμάνσεις που δεν φαίνονται στα πραγματικά δεδομένα. Από την άλλη η πρόβλεψη του δείκτη PoA δεν ακολουθεί ούτε την τάση των παρατηρήσεων ούτε τις διακυμάνσεις αυτών. Αυτό πιθανώς σημαίνει ότι οι τεχνολογίες FBA και PoA δεν επηρεάζονται τόσο πολύ από τις υπόλοιπες και ακολουθούν την δική τους αυτόνομη συμπεριφορά και έτσι οι προβλέψεις με βάση τις άλλες τεχνολογίες δυσκολεύονται να προβλέψουν τις τιμές τους. Αυτό ως ένα σημείο επαληθεύει και την υπόθεση της τάσης που οφείλεται στην δημοτικότητα στον περίγυρο των επενδυτών. Τα πιο δημοφιλή κρυπτονομίσματα είναι τύπου PoW και PoS, και με βάση την προηγούμενη υπόθεση κάποιος θα περίμενε να έχουν αρκετά καλή γραμμική συσχέτιση, η οποία ως έναν βαθμό παρατηρείται, ενώ τα κρυπτονομίσματα FBA και PoA, δεν είναι τόσο δημοφιλή και θεωρητικά τουλάχιστον δεν έχουν την ίδια αύξηση που θα προέκυπτε από δημοτικότητα των άλλων τεχνολογιών. Ως έναν βαθμό λοιπόν, φαίνεται πως οι πιο αφανείς τεχνολογίες έχουν μικρότερη γραμμική συσχέτιση. Παρόλα αυτά, τουλάχιστον για τον FBA ακολουθείται μία ανοδική τάση ή οποία μπορεί να εξηγηθεί στην αύξηση της γενικής δημοτικότητας των κρυπτονομισμάτων.

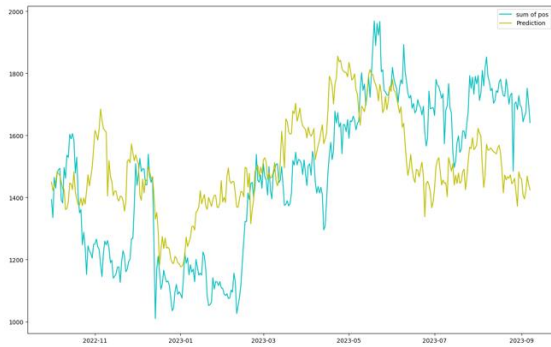
6.2.2 Μοντέλο με χρήση βέλτιστων ολισθήσεων και ολισθήσεων κατά μία ημέρα

Το επόμενο μοντέλο που θα χρησιμοποιηθεί βελτιώνει το πρωταρχικό, προσθέτοντας στην είσοδο τα δεδομένα των δεικτών μετατοπισμένα κατά μία μέρα όπου αυτά δεν υπάρχουν. Έτσι έχοντας πάλι ως παράδειγμα την πρόβλεψη του PoS, οι δείκτες PoW και FBA θα προστεθούν σαν δεδομένα στην είσοδο του μοντέλου. Ο δείκτης PoA είναι ήδη μετατοπισμένος κατά μία ημέρα οπότε δεν χρειάζεται να προστεθεί δεύτερη φορά.

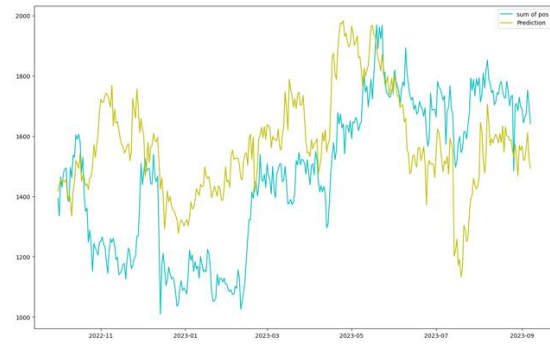
Πρωτόκολλο	RMSE	MAPE
PoW	3924.86	0.174424
PoS	251.614	0.159222
FBA	0.0939303	0.225101
PoA	80.8711	0.267591

Πίνακας 6. Αποτελέσματα δεικτών ακρίβειας μοντέλου βέλτιστων ολισθήσεων και ολισθήσεων κατά μία ημέρα

Παρατηρούμε από τους δείκτες ακρίβειας ότι αρχικά τα σφάλματα του δείκτη PoW προφανώς δεν αλλάζουν. Αυτό είναι λογικό γιατί τα βέλτιστα διαστήματα ολίσθησης ήταν ήδη η μία ημέρα και έτσι το μοντέλο εδώ είναι το ίδιο. Από την άλλη παρατηρούμε ότι για τον δείκτη PoS παρατηρούμε αύξηση στα σφάλματα, ενώ για τους δείκτες FBA και PoA μείωση αυτών.



Εικόνα 33. Πρόβλεψη του δείκτη PoS με την χρήση του μοντέλου 6.2.1

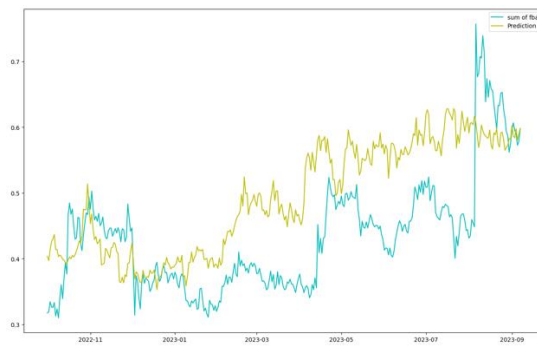


Εικόνα 34. Πρόβλεψη του δείκτη PoS με την χρήση του μοντέλου 6.2.2

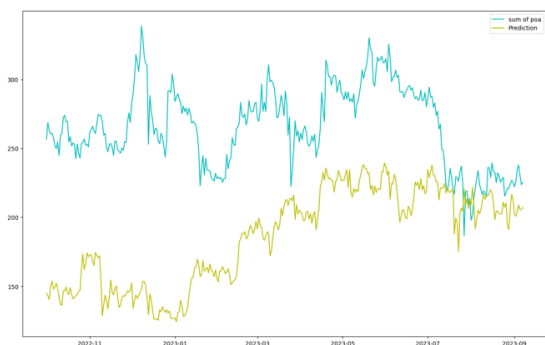
Συγκρίνοντας την γραφική απεικόνιση των δύο μοντέλων για τον PoS παρατηρούμε ότι το δεύτερο μοντέλο συμπεριφέρεται καλύτερα στο κομμάτι της αναπαράστασης της γραμμικής συσχέτισης παρόλα που παρουσιάζει μεγαλύτερα σφάλματα. Για παράδειγμα τον Ιούλιο και Αύγουστο του 2023, η πρόβλεψη στο δεύτερο μοντέλο ακολουθεί την απότομη πτώση της τιμής των παρατηρήσεων, ενώ στο πρώτο μοντέλο αυτό δεν γίνεται. Παρόλα αυτά οι πτώσεις των προβλέψεων είναι πιο απότομες και αυτός είναι ο κύριος λόγος που το μοντέλο έχει μεγαλύτερα σφάλματα. Η μεγαλύτερη πτώση αυτή μπορεί να οφείλεται στην επιρροή από την υψηλή τιμή των κρυπτονομισμάτων PoW, που σε μεγάλες πτώσεις λόγω της μεγαλύτερης τιμής τους, επηρεάζει πολύ πιο απότομα την πρόβλεψη του PoS. Επίσης, τα επιπλέον δεδομένα των κοντινών ημερών, τουλάχιστον σε απότομες πτώσεις ή αυξήσεις μπορεί να είναι προϊόν ασυνήθιστων γεγονότων και έτσι επηρεάζουν περισσότερο το μοντέλο και είναι πιθανώς ένας ακόμα λόγος που παρατηρούμε καλύτερη γραμμική συσχέτιση. Παρόμοιες συμπεριφορές αύξησης της γραμμικής συσχέτισης παρατηρούνται και συγκρίνοντας τις γραφικές απεικονίσεις για τα FBA και PoA. Έτσι σε γενικές γραμμές το δεύτερο μοντέλο είναι ελάχιστα καλύτερο από το πρώτο.



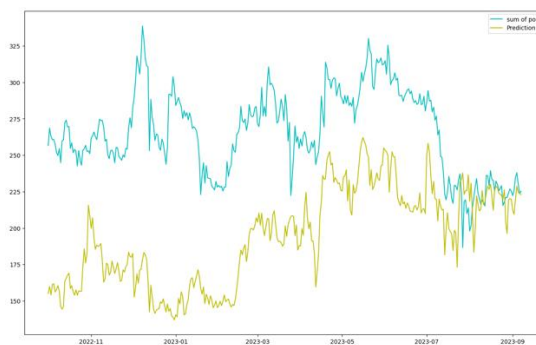
Εικόνα 35. Πρόβλεψη του δείκτη FBA με την χρήση του μοντέλου 6.2.1



Εικόνα 36. Πρόβλεψη του δείκτη FBA με την χρήση του μοντέλου 6.2.2



Εικόνα 37. Πρόβλεψη του δείκτη PoA με την χρήση του μοντέλου 6.2.1



Εικόνα 38. Πρόβλεψη του δείκτη PoA με την χρήση του μοντέλου 6.2.2

6.2.3 Μοντέλο με χρήση βέλτιστων ολισθήσεων, ολισθήσεων κατά μία ημέρα και ολισθήσεων γειτονικών όρων επιλεγμένων χειροκίνητα (cherry-picked)

Το συγκεκριμένο μοντέλο προσπαθεί να εξηγήσει, την υπόθεση του αν περισσότερα δεδομένα δίνουν καλύτερα αποτελέσματα. Έτσι χτίζοντας στο δεύτερο μοντέλο προστίθενται και επιπλέον δείκτες μετατοπισμένοι κατά x ημέρες κοντά σε αυτές που παρουσιάζουν την καλύτερη γραμμική συσχέτιση. Οι ημέρες μετατόπισης προτιμήθηκε να επιλεγθούν χειροκίνητα και όχι αλγοριθμικά (πχ. Οι ημέρες με τον δεύτερο καλύτερο συντελεστή συσχέτισης), για διεύρυνση των δεδομένων. Για παράδειγμα, σε περιπτώσεις όπου παρουσιαζόταν καλή γραμμική συσχέτιση κάποιες μέρες μακριά από την καλύτερη μετατόπιση προτιμήθηκαν αυτές, σε σχέση με την διπλανή ημέρα στην καλύτερη, για επιπρόσθετη διαφοροποίηση. Παρόλα αυτά ο γνώμονας επιλογής ήταν πάντα η βελτίωση του μοντέλου και έτσι επιλέχθηκαν μέρες μετατόπισης με πολύ κοντινούς συντελεστές συσχέτισης στον δεύτερο καλύτερο.

Πρωτόκολλο	RMSE	MAPE
PoW	4004.12	0.179081
PoS	254.842	0.159699
FBA	0.0943778	0.225101
PoA	80.906	0.267982

Πίνακας 7. Αποτελέσματα δεικτών ακρίβειας μοντέλου βέλτιστων ολισθήσεων, ολισθήσεων κατά μία ημέρα και cherry-picked ημερών

Συγκρίνοντας τα αποτελέσματα του δεύτερου και τρίτου μοντέλου, παρατηρούμε ότι οι δείκτες ακρίβειας έχουν όλοι αυξηθεί δείχνοντας αύξηση των σφαλμάτων σε όλες τις περιπτώσεις. Οι αλλαγές αυτές είναι ελάχιστες σε όλες τις χρονοσειρές, χωρίς να ξεπερνάνε το 1%, συγκρίνοντας τους δείκτες MAPE. Παρόλα αυτά, το συγκεκριμένο μοντέλο αναδεικνύει ότι τα περισσότερα δεδομένα δεν σημαίνουν απαραίτητα καλύτερη πρόβλεψη. Το δεύτερο μοντέλο χρησιμοποιώντας τις καλύτερες ημέρες, και αυτές της μίας ημέρας για καλύτερη αναπαράσταση των απότομων αλλαγών έχει αρκετά δεδομένα για να παράξει ικανοποιητικές προβλέψεις και δεδομένα γειτονικών ημερών απλά παράγουν θόρυβο. Οι γραφικές απεικονίσεις σε αυτό το στάδιο παραλείπονται, γιατί είναι σημαντικά μικρές οι αλλαγές στην συμπεριφορά των μοντέλων, ειδικά στα μικρά μεγέθη των FBA και PoA, όπου οι απεικονίσεις φαίνονται σχεδόν πανομοιότυπες με αυτές του δεύτερου μοντέλου.

6.2.4 Μοντέλο καλύτερων ολισθήσεων με επανεκπαίδευση

Μια άλλη βελτίωση που μπορεί να γίνει στο μοντέλο, είναι η εισαγωγή των πιο πρόσφατων δεδομένων στα δεδομένα εκπαίδευσης. Εξαιτίας της απότομης μεταβολής του επιπέδου των παρατηρήσεων των κρυπτονομισμάτων, λόγω της αστάθειας που παρουσιάζουν, θεωρείται σημαντικό να προμηθευτεί το μοντέλο με όσο το δυνατόν νεότερα δεδομένα γίνεται. Το πρόβλημα εδώ είναι ο χωρισμός των δεδομένων, και η ανάγκη ύπαρξης ενός συνόλου για την αξιολόγηση αυτού. Το πρόβλημα αυτό λύνεται με μία διαδικασία που ονομάζεται επανεκπαίδευση (retrain).

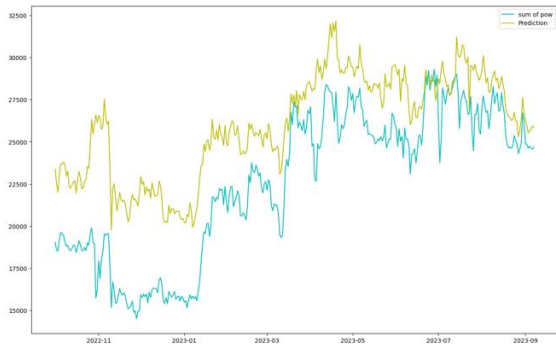
Η επανεκπαίδευση παίρνει δεδομένα που ήδη χρησιμοποιήθηκαν στην αξιολόγηση του και τα προσθέτει στα δεδομένα εκπαίδευσης. Η διαδικασία αυτή γίνεται σε στάδια, όπου το μοντέλο αξιολογείται για κάποια διαστήματα του συνόλου αξιολόγησης, και μετά το διάστημα των δεδομένων που χρησιμοποιήθηκε προστίθεται στα δεδομένα εκπαίδευσης. Η λύση αυτή μας παρέχει τα πιο πρόσφατα δεδομένα για την προσπάθεια εξάλειψης των προβλημάτων μακροπρόθεσμης πρόβλεψης και της απότομης αλλαγής του επιπέδου των χρονοσειρών. Στην περίπτωση του μοντέλου, το διάστημα επανεκπαίδευσης που χρησιμοποιείται είναι οι 30 ημέρες. Τα δεδομένα εισόδου του μοντέλου επανεκπαίδευσης είναι τα ίδια με αυτά του δεύτερου μοντέλου που αναφέρθηκε στην ενότητα 6.2.2.

Πρωτόκολλο	RMSE	MAPE	Σωστή πρόβλεψη τάσης (%)
PoW	3097.61	0.129014	76.9
PoS	219.027	0.131163	58.77
FBA	0.0878038	0.180388	59.94
PoA	70.3363	0.223313	60.81

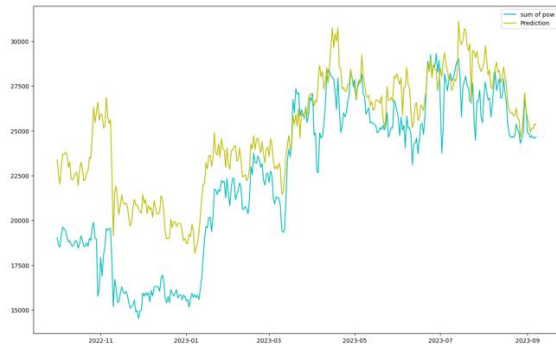
Πίνακας 8. Αποτελέσματα δεικτών ακρίβειας μοντέλου επανεκπαίδευσης

Από τα αποτελέσματα των δεικτών ακρίβειας του μοντέλου, παρατηρείται σημαντική μείωση του ποσοστού των σφαλμάτων σε όλα τα πρωτόκολλα, βλέποντας μείωση σχεδόν 5% στα σφάλματα στην καλύτερη περίπτωση συγκρίνοντας τους δείκτες MAPE του αλγορίθμου PoW. Στον πίνακα σε αυτό το στάδιο προστίθεται και ένας δείκτης σωστής πρόβλεψης τάσης ο οποίος είναι ένας δείκτης που ορίζεται από εμάς, ως το ποσοστό των ημερήσιων δεδομένων που το μοντέλο προβλέπει την σωστή αύξηση ή μείωση τους σε σχέση με την τιμή της προηγούμενης ημέρας. Ο δείκτης αυτός είναι πολύ σημαντικός για προβλέψεις στον χρηματιστηριακό τομέα, όπου ένα ποσοστό 100% θα σήμαινε κέρδος σε κάθε πρόβλεψη του μοντέλου, ανεξαρτήτως του αν το μοντέλο μπορεί να προβλέψει με ακρίβεια ή όχι την τελική τιμή της μετοχής. Ο συγκεκριμένος δείκτης είναι και μία μορφή απεικόνισης αριθμητικά της γραμμικής συσχέτισης της πρόβλεψης με την πραγματική τιμή. Ουσιαστικά τα αποτελέσματα του δείκτη δείχνουν ότι η πρόβλεψη για τον αλγόριθμο PoW έχει την καλύτερη γραμμική συσχέτιση με τα πραγματικά δεδομένα σε αρκετά ικανοποιητικό βαθμό. Επίσης ο συγκεκριμένος δείκτης τονίζει ότι ο δείκτης MAPE δεν είναι απαραίτητα ο καλύτερος δείκτης για την σύγκριση που προσπαθεί να επιτευχθεί και επαναλαμβάνει κάποια συμπεράσματα που αναφέρθηκαν στην ενότητα 6.2.1 γραφικά, όπου τονίστηκε ότι ο PoW έχει την καλύτερη γραμμική συσχέτιση παρόλης της τιμής του MAPE. Τέλος

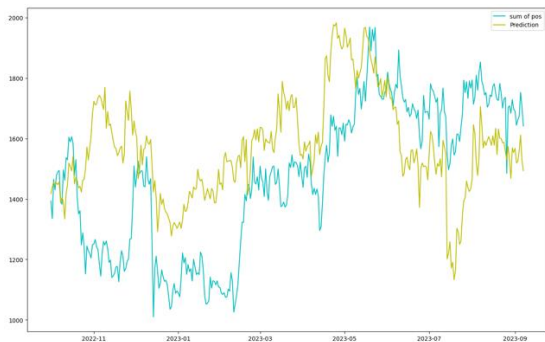
παρόλο που για τα άλλα πρωτόκολλα οι δείκτες απέχουν περίπου 15% από τον δείκτη του PoW, τα αποτελέσματα είναι αρκετά ενθαρρυντικά, γιατί σε μία μετοχή χωρίς μεγάλες αυξομειώσεις που κινείται παρόμοια σε πτώσεις και αυξήσεις, κάθε δείκτης πάνω του 50% προδίδει κέρδος.



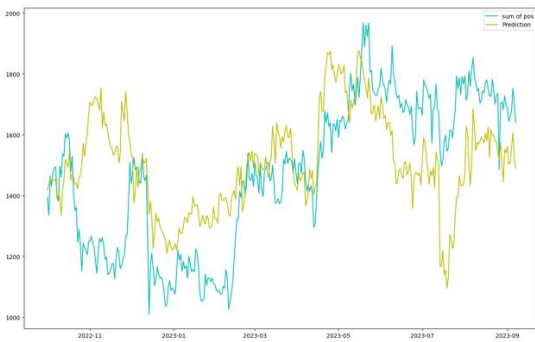
Εικόνα 39. Πρόβλεψη του δείκτη PoW με την χρήση του μοντέλου 6.2.2



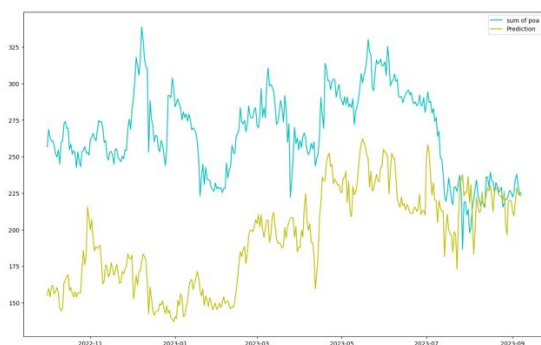
Εικόνα 40. Πρόβλεψη του δείκτη PoW με την χρήση του μοντέλου 6.2.4



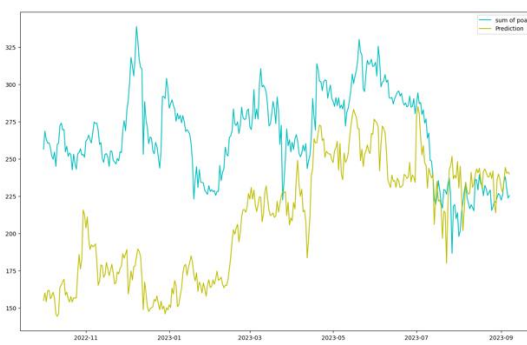
Εικόνα 41. Πρόβλεψη του δείκτη PoS με την χρήση του μοντέλου 6.2.2



Εικόνα 42. Πρόβλεψη του δείκτη PoS με την χρήση του μοντέλου 6.2.4



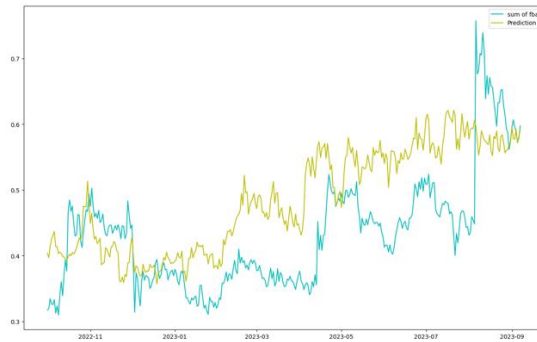
Εικόνα 43. Πρόβλεψη του δείκτη PoA με την χρήση του μοντέλου 6.2.2



Εικόνα 44. Πρόβλεψη του δείκτη PoA με την χρήση του μοντέλου 6.2.4



Εικόνα 45. Πρόβλεψη του δείκτη FBA με την χρήση του μοντέλου 6.2.2



Εικόνα 46. Πρόβλεψη του δείκτη FBA με την χρήση του μοντέλου 6.2.4

Από τις παραπάνω γραφικές απεικονίσεις παρατηρείται ότι η συμπεριφορά των θετικών/αρνητικών προβλέψεων δεν επηρεάζεται τόσο καθώς υπάρχουν πολλές ομοιότητες στις προβλέψεις των δύο μοντέλων. Παρόλα αυτά, αυτό που αλλάζει στο μοντέλο του 6.2.4 είναι η ακρίβεια των προβλέψεων όπου το επίπεδο έχει πλησιάσει σε πολλές περιπτώσεις πιο κοντά στα πραγματικά δεδομένα. Αυτό φαίνεται σε πολλά σημεία στις προβλέψεις όλων των αλγορίθμων.

6.2.5 Μοντέλο καλύτερων ολισθήσεων με επανεκπαίδευση και αφαίρεση των παλαιότερων δεδομένων

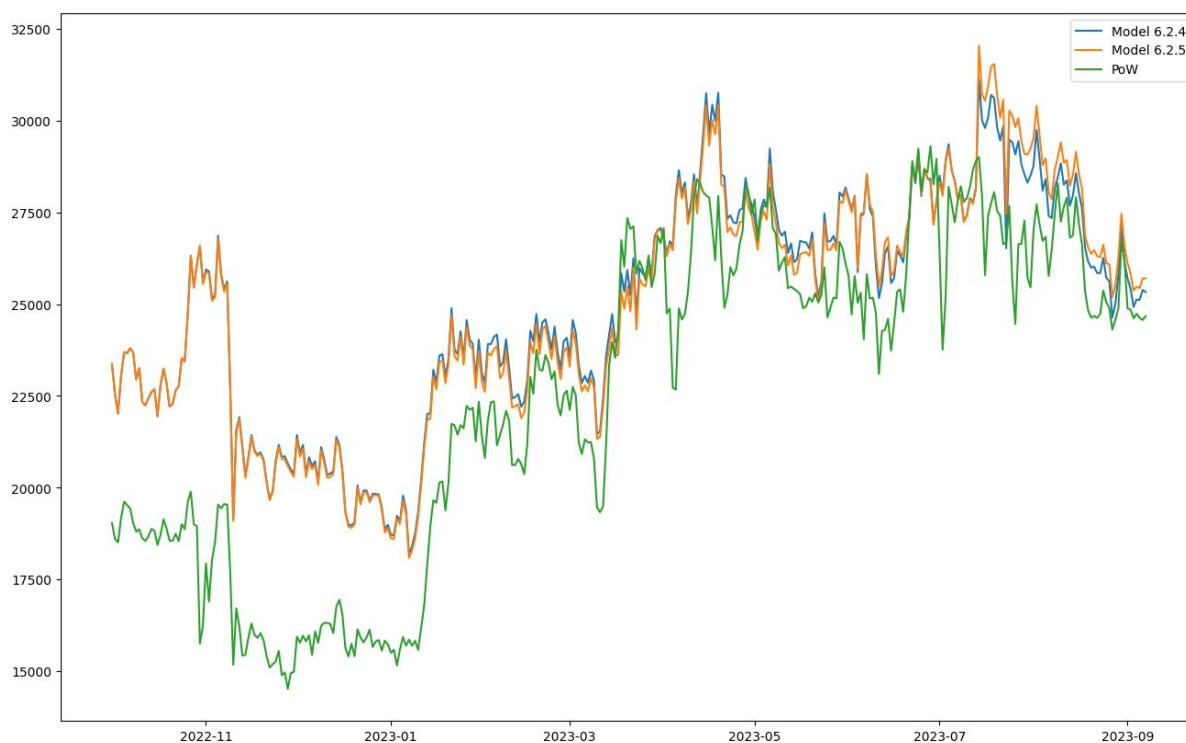
Για την επιπρόσθετη βελτίωση του προηγούμενου μοντέλου θα ελεγχθεί η υπόθεση της επιρροής των παλαιότερων δεδομένων στο μοντέλο πρόβλεψης. Η συγκεκριμένη υπόθεση βασίζεται στις ίδιες αρχές με αυτές του μοντέλου της μη γραμμικής τάσης όπου τα παλαιότερα δεδομένα θεωρείται ότι επηρεάζουν λιγότερο την πιο πρόσφατη τιμή. Στην περίπτωση του μοντέλου, απλοποιείται η υπόθεση αυτή, αφαιρώντας τα παλαιότερα δεδομένα καθώς αλλάζουν τα διαστήματα επανεκπαίδευσης. Έτσι στο μοντέλο, όταν προστίθενται 30 ημέρες στα δεδομένα εκπαίδευσης, θα αφαιρούνται οι 30 παλαιότερες ημέρες. Το αποτέλεσμα που προκύπτει είναι το εξής.

Πρωτόκολλο	RMSE	MAPE	Σωστή πρόβλεψη τάσης (%)
PoW	3097.93	0.128822	77.48
PoS	225.634	0.134878	59.35
FBA	0.0840457	0.172429	59.64
PoA	70.1768	0.222619	60.52

Πίνακας 9. Αποτελέσματα δεικτών ακρίβειας μοντέλου επανεκπαίδευσης 6.2.5

Όσον αφορά τους δείκτες MAPE όλοι μειώνονται ελάχιστα, εκτός από αυτό του PoS, που αυξάνεται αλλά σε πολύ μικρή κλίμακα. Για τους δείκτες σωστής πρόβλεψης έχουμε παρόμοια συμπεριφορά όπου έχουμε μικρή πτώση ή αύξηση του δείκτη, αλλά σε αμελητέο βαθμό. Παρόλα αυτά για τους δείκτες πρόβλεψης οι αυξήσεις είναι της τάξης του 1% ενώ οι μειώσεις του 0.5%, οπότε θεωρείται ελάχιστα καλύτερο το παρόν μοντέλο. Παρόλα αυτά, τόσο γραφικά όσο και αριθμητικά, τα δύο μοντέλα (6.2.4 και 6.2.5) έχουν ελάχιστες διαφορές μεταξύ τους. Συγκρίνοντας τις διαφοροποιήσεις των δύο μοντέλων, παρατηρείται ότι το μοντέλο του 6.2.5 τις περισσότερες φορές είναι πιο κοντά στα πραγματικά δεδομένα. Για παράδειγμα, για το χρονικό διάστημα από 01-

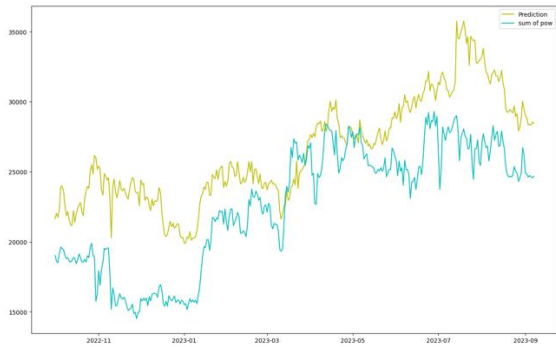
2023 έως 06-2023, η προβλέψεις του μοντέλου 6.2.5 επηρεάζονται λιγότερο από τις αυξημένες τιμές που παρατηρήθηκαν τον 11-2022 και οι προβλέψεις ακολουθούν περισσότερο τις τιμές του πιο πρόσφατου μειωμένου επιπέδου (01-2023). Άρα ως έναν βαθμό επιτυγχάνεται ο στόχος με την αφαίρεση παλαιότερων δεδομένων. Παρόλο που τα μοντέλα 6.2.4 και 6.2.5 έχουν μικρές διαφορές μεταξύ τους, στα πλαίσια των επόμενων ενοτήτων θα θεωρείται σαν βέλτιστο το μοντέλο του 6.2.5, και πάνω σε αυτό θα προστεθούν οι βασικοί χρηματιστηριακοί δείκτες του χρυσού, του αργού πετρελαίου και του S&P500.



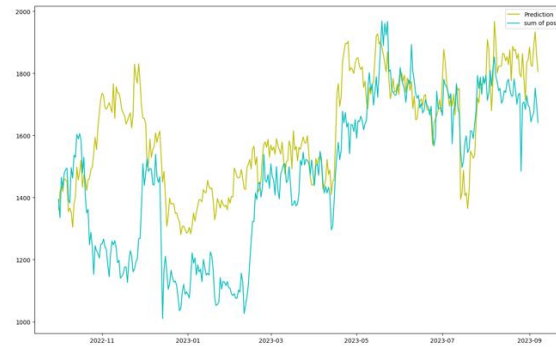
Εικόνα 47. Διαφορά των μοντέλων 6.2.4 και 6.2.5 για την πρόβλεψη του PoW

6.2.6 Βέλτιστο μοντέλο με προσθήκη του δείκτη S&P500

Στο βέλτιστο μοντέλο που παρουσιάστηκε παραπάνω θα προστεθεί ο δείκτης S&P500 για την μελέτη της επιρροής του και της σχέσης που αυτός έχει με τα κρυπτονομίσματα. Έτσι το μοντέλο που μελετάται τώρα θα έχει τις τρεις χρονοσειρές μετατοπισμένες με τον τρόπο που μελετήθηκε παραπάνω και του δείκτη S&P500 μετατοπισμένο κατά μία ημέρα. Στο σημείο αυτό επιλέχθηκε η ολίσθηση κατά μία ημέρα τόσο για τον S&P500 όσο και για τους άλλους δείκτες που θα μελετηθούν παρακάτω, γιατί μελετώντας την εικόνα 22, είναι εύκολα κατανοητό ότι δεν υπάρχει μεγάλη γραμμική συσχέτιση μεταξύ των δεικτών και των κρυπτονομισμάτων. Παρόλα αυτά, κάποιες ακραίες συνθήκες της αγοράς (πχ. Η απότομη πτώση όλων των δεικτών στις αρχές του 2020), που πιθανώς δημιουργούνται από εξωγενείς απρόβλεπτους παράγοντες, επηρεάζουν τόσο τα κρυπτονομίσματα όσο και τους δείκτες, οπότε η ολίσθηση μίας ημέρας μας δίνει τα καλύτερα αποτελέσματα σε αυτές τις περιπτώσεις. Τα αποτελέσματα του επαυξημένου μοντέλου είναι τα εξής.



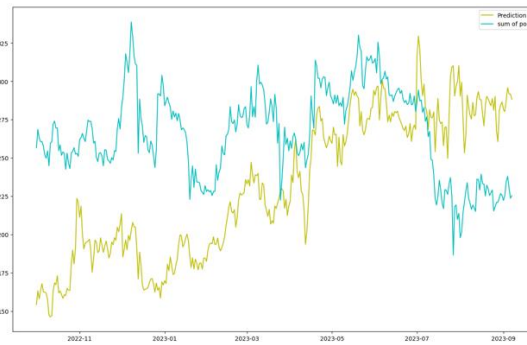
Εικόνα 48. Πρόβλεψη του δείκτη PoW με την χρήση του μοντέλου 6.2.6



Εικόνα 49. Πρόβλεψη του δείκτη PoS με την χρήση του μοντέλου 6.2.6



Εικόνα 50. Πρόβλεψη του δείκτη FBA με την χρήση του μοντέλου 6.2.6



Εικόνα 51. Πρόβλεψη του δείκτη PoA με την χρήση του μοντέλου 6.2.6

Πρωτόκολλο	RMSE	MAPE	Σωστή πρόβλεψη τάσης (%)
PoW	4604.49	0.198651	72.8
PoS	219.042	0.127685	57.01
FBA	0.0827438	0.159211	57.89
PoA	63.4665	0.212666	63.74

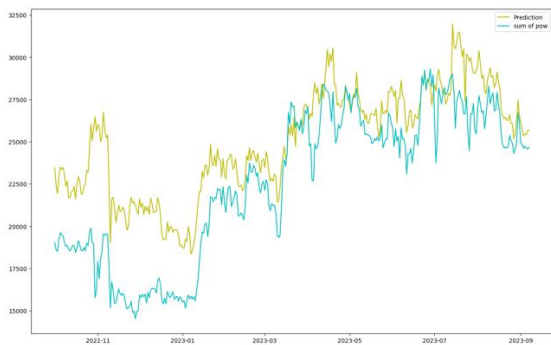
Πίνακας 10. Αποτελέσματα δεικτών ακρίβειας μοντέλου με προσθήκη του S&P500

Παρατηρώντας τις γραφικές παραστάσεις, βλέπουμε ότι ο δείκτης S&P500 λειτουργεί ως ένας παράγοντας εξομάλυνσης των προβλέψεων. Συγκρίνοντας τις προβλέψεις των μοντέλων 6.2.5 και 6.2.6 βλέπουμε ότι ο δείκτης S&P500 μειώνει την γραμμική συσχέτιση που υπάρχει σε πολλά σημεία για όλα τα πρωτόκολλα, αλλά από την άλλη εξομαλύνει το επίπεδο των προβλέψεων έτσι ώστε αυτό να τείνει πιο κοντά σε αυτό των πραγματικών δεδομένων. Το γεγονός αυτό φαίνεται και στον πίνακα των σφαλμάτων ακρίβειας, όπου σε πολλές περιπτώσεις ο δείκτης MAPE δίνει καλύτερα αποτελέσματα, αλλά ο δείκτης σωστής πρόβλεψης χειρότερα σε σχέση με το μοντέλο 6.2.5. Παρόλα αυτά ένα κύριο μειονέκτημα της προσθήκης του δείκτη στο μοντέλο είναι η επιρροή του σε προσθήκη έντονης αυξητικής τάσης στις τελικές προβλέψεις. Αυτό φαίνεται έντονα στην εικόνα 51 όπου οι πραγματικές τιμές του PoA έχουν ένα στάσιμο επίπεδο, και οι προβλέψεις λόγω της προσθήκης του δείκτη S&P έχουν έντονη αυξητική τάση, μειώνοντας σημαντικά την γραμμική συσχέτιση των προβλέψεων με

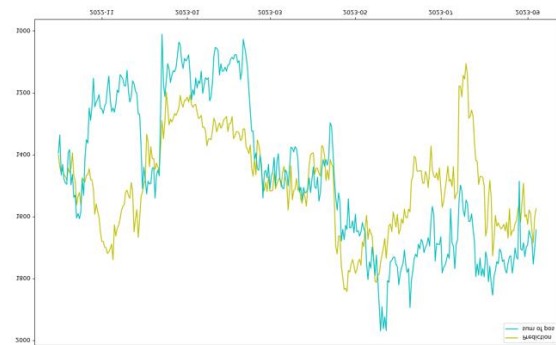
τα δεδομένα. Γενικά παρατηρείται ότι ο δείκτης εξομαλύνει τις προβλέψεις και ως αποτέλεσμα έχει την μείωση της μεγάλης αστάθειας στις προβλέψεις και οι προβλέψεις να τείνουν κοντινότερα στα πραγματικά δεδομένα στις περισσότερες των περιπτώσεων. Παρόλα αυτά, τα κρυπτονομίσματα έχουν ασταθή φύση και χάνεται ένας σημαντικός παράγοντας γραμμικής συσχέτισης με την προσθήκη του S&P κάνοντας την συμπεριφορά του μοντέλου χειρότερη με βάση τους στόχους για σωστή πρόβλεψη της τάσης πράγμα που επαληθεύεται και από το γεγονός ότι οι δείκτες σωστής πρόβλεψης είναι χειρότεροι για τις τρεις από τις τέσσερις χρονοσειρές που μελετώνται.

6.2.7 Βέλτιστο μοντέλο με προσθήκη του δείκτη του αργού πετρελαίου

Η προσθήκη του δείκτη του αργού πετρελαίου γίνεται με την ίδια λογική όπως και για τον δείκτη S&P500. Τα αποτελέσματα του τελικού μοντέλου είναι τα εξής.



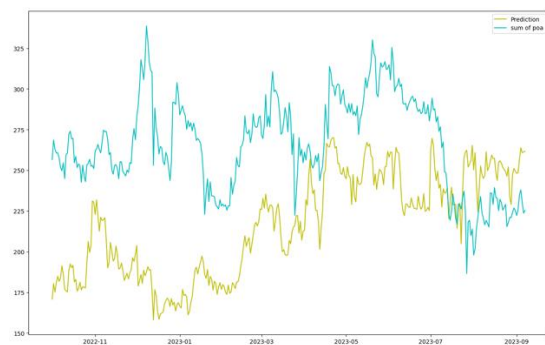
Εικόνα 52. Πρόβλεψη του δείκτη PoW με την χρήση του μοντέλου 6.2.7



Εικόνα 53. Πρόβλεψη του δείκτη PoS με την χρήση του μοντέλου 6.2.7



Εικόνα 54. Πρόβλεψη του δείκτη FBA με την χρήση του μοντέλου 6.2.7



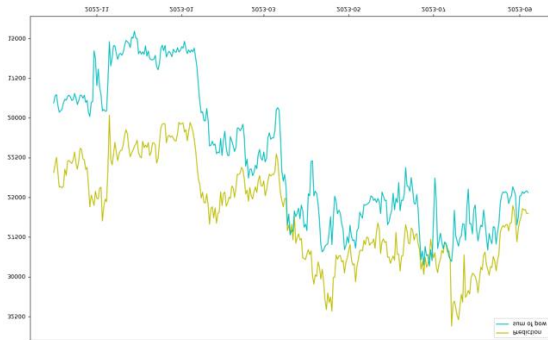
Εικόνα 55. Πρόβλεψη του δείκτη PoA με την χρήση του μοντέλου 6.2.7

Πρωτόκολλο	RMSE	MAPE	Σωστή πρόβλεψη τάσης (%)
PoW	3155.92	0.132319	76.6
PoS	206.063	0.119729	59.64
FBA	0.08794	0.182799	60.52
PoA	63.843	0.209985	59.35

Συγκρίνοντας τα σφάλματα ακρίβειας των μοντέλων 6.2.6 και 6.2.7, παρατηρείται ότι ο δείκτης του αργού πετρελαίου δίνει τις περισσότερες των περιπτώσεων καλύτερα αποτελέσματα τόσο στον δείκτη MAPE όσο και στον δείκτη σωστής πρόβλεψης. Συγκρίνοντας τώρα τα μοντέλα 6.2.5 και 6.2.7, τα αποτελέσματα δεν έχουν μεγάλες διαφορές και κάποιες φορές είναι το ένα μοντέλο καλύτερο ενώ άλλες φορές είναι το άλλο. Αυτό πιθανώς δείχνει ότι ο δείκτης του αργού πετρελαίου έχει καλύτερη γραμμική σχέση με τα κρυπτονομίσματα σε σχέση με τον δείκτη S&P. Αυτό φαίνεται και γραφικά, όπου με την εφαρμογή του δείκτη του πετρελαίου δεν αλλοιώνεται τόσο η γραμμική συσχέτιση της πρόβλεψης με τα τελικά αποτελέσματα όπως αλλοιωνόταν στο μοντέλο 6.2.6. Έτσι ο δείκτης του πετρελαίου παρέχει τα πλεονεκτήματα της εξομάλυνσης που παρείχε και ο δείκτης S&P, χωρίς να μειώνει την γραμμική συσχέτιση που παρουσιάζουν τα κρυπτονομίσματα κάνοντας τον πιθανώς καλό παράγοντα διαφοροποίησης για το τελικό μοντέλο.

6.2.8 Βέλτιστο μοντέλο με προσθήκη του δείκτη του χρυσού

Η προσθήκη του δείκτη του χρυσού γίνεται με την ίδια λογική όπως και για τους παραπάνω δείκτες. Τα αποτελέσματα του τελικού μοντέλου είναι τα εξής.



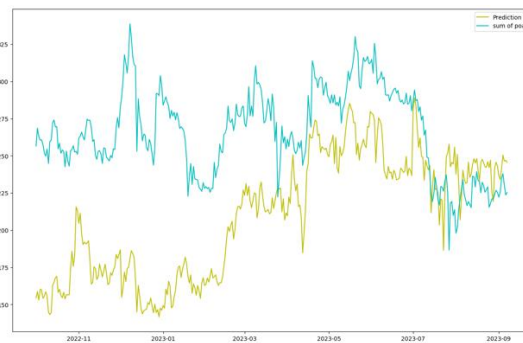
Εικόνα 56. Πρόβλεψη του δείκτη PoW με την χρήση του μοντέλου 6.2.8



Εικόνα 57. Πρόβλεψη του δείκτη PoS με την χρήση του μοντέλου 6.2.8



Εικόνα 58. Πρόβλεψη του δείκτη FBA με την χρήση του μοντέλου 6.2.8



Εικόνα 59. Πρόβλεψη του δείκτη PoA με την χρήση του μοντέλου 6.2.8

Πρωτόκολλο	RMSE	MAPE	Σωστή πρόβλεψη τάσης (%)
PoW	3917.17	0.173685	75.43
PoS	218.67	0.131379	58.77
FBA	0.0713117	0.135309	58.77
PoA	71.2864	0.226304	60.52

Πίνακας 12. Αποτελέσματα δεικτών ακρίβειας μοντέλου με προσθήκη του χρυσού

Με την προσθήκη του χρυσού τα αποτελέσματα πάλι συγκρατούν έναν μεγάλο βαθμό γραμμικής συσχέτισης και παράλληλα εξομαλύνονται. Ο δείκτης του χρυσού παρουσιάζει και αυτός καλύτερα αποτελέσματα από τον S&P στις περισσότερες των περιπτώσεων τόσο στον δείκτη MAPE όσο και στον δείκτη σωστής πρόβλεψης. Από την άλλη, σε σύγκριση με τα ποσοστά σωστής πρόβλεψης τόσο του 6.2.5 όσο και του μοντέλου 6.2.7, η προσθήκη του χρυσού χειροτερεύει τις προβλέψεις. Οπότε παρόλο που τα αποτελέσματα του δείκτη MAPE είναι ενθαρρυντικά, τα αποτελέσματα πρόβλεψης δεν βελτιώνονται. Έτσι όπως και με την προσθήκη του δείκτη S&P500, έτσι και με την προσθήκη του χρυσού, δεν λαμβάνονται ικανοποιητικά αποτελέσματα.

6.3 Αξιολόγηση αποτελεσμάτων και πρόβλεψη τιμής

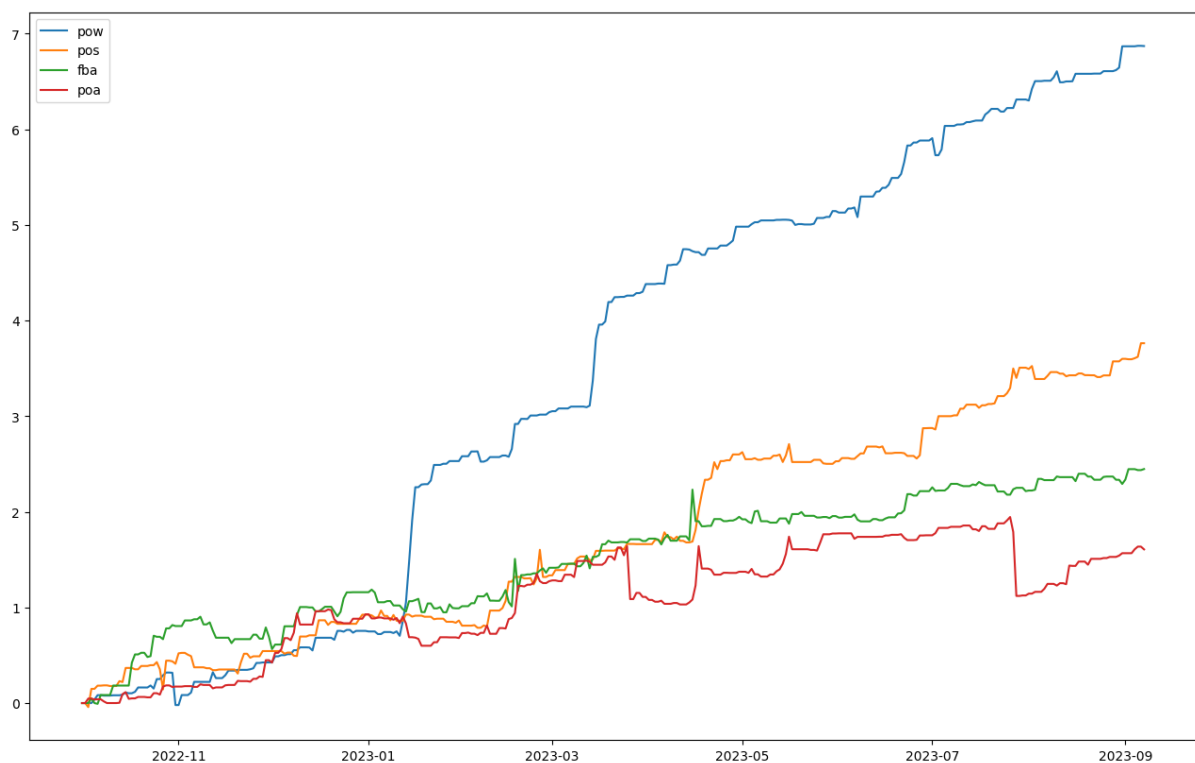
Από τα παραπάνω αποτελέσματα παρατηρούμε ότι το καλύτερο μοντέλο είναι αυτό με την προσθήκη επανεκπαίδευσης και χρήση τόσο των καλύτερων ολισθήσεων όσο και των ολισθήσεων κατά μία ημέρα. Οι δείκτες MAPE επειδή είναι της τάξης του 10-20% είναι πιθανώς απαγορευτικοί για πρόβλεψη της ακριβούς τιμής, αλλά αυτό οφείλεται και στην χρήση ενός απλού μοντέλου όπως αυτό της γραμμικής παλινδρόμησης. Παρόλα αυτά, οι δείκτες σωστής πρόβλεψης είναι αρκετά ενθαρρυντικοί καθώς υπονοούν ότι το μοντέλο μπορεί να προβλέψει με αρκετά καλά ποσοστά, ειδικά στο μοντέλο PoW, την κίνηση της χρονοσειράς της κάθε τεχνολογίας. Για καλύτερη αξιολόγηση του μοντέλου σε χρηματιστηριακές προβλέψεις θα το χρησιμοποιήσουμε εικονικά για την λήψη επενδυτικών αποφάσεων στο σύνολο των δεδομένων αξιολόγησης.

Το μοντέλο που θα χρησιμοποιήσουμε είναι το βέλτιστο μοντέλο που αναφέρθηκε και παραπάνω και αυτό επενδύει 1€ στην κάθε εικονική μετοχή, κάθε φορά που το μοντέλο προβλέπει αύξηση, και πουλάει όλες τις διαθέσιμες μετοχές όταν το μοντέλο προβλέπει μείωση. Στον πραγματικό κόσμο οι τιμές της κάθε μετοχής κυμαίνονται συνεχώς, και η αγορά μετοχών έχει ένα κομμάτι προμήθειας από τον μεσίτη που αγοράζει την μετοχή, αλλά για απλοποίηση στα πλαίσια της εργασίας, οι μετοχές είναι σαν να αγοράζονται στην τιμή κλεισίματος του χρηματιστηρίου κάθε μέρα χωρίς προμήθεια. Η σύγκριση του μοντέλου μας θα γίνει με βάση δύο απλά μοντέλα τα οποία βασίζονται στην απλούστερη μορφή επένδυσης, της αγοράς μίας μετοχής σε χαμηλή τιμή και πώληση της σε υψηλή (buy & hold). Το πρώτο μοντέλο θα αγοράζει την μετοχή στην αρχή των δεδομένων αξιολόγησης και θα την πουλάει στο τέλος των δεδομένων αξιολόγησης, δίνοντας το πιθανό κέρδος σε περίπτωση επένδυσης χωρίς κάποιο κριτήριο. Το δεύτερο μοντέλο ακολουθεί την ίδια λογική επένδυσης, αγοράζοντας όμως στην χαμηλότερη δυνατή τιμή και πουλώντας στην υψηλότερη, δίνοντας το μέγιστο πιθανό κέρδος της λογικής buy & hold.

Πρωτόκολλο	Κέρδος Μοντέλου	Κέρδος αφελούς buy & hold	Κέρδος βέλτιστου buy & hold
PoW	6.88012	1.29638	2.01831
PoS	3.76336	1.17783	1.94915
FBA	2.51283	1.88159	2.44263
PoA	1.6072	0.878158	1.48448

Πίνακας 13. Σύγκριση αποτελεσμάτων επένδυσης του βέλτιστου μοντέλου

Τα αποτελέσματα του πίνακα είναι πρακτικά το πόσο πολλαπλασιάστηκε η μονάδα επένδυσης μας, έτσι για το παράδειγμα, στο μοντέλο PoW το 1€ γίνεται 6.88€ από το μοντέλο, ενώ αγοράζοντας και πολώνοντας την καλύτερη δυνατή στιγμή γίνεται 2.08€. Σε όλες τις περιπτώσεις βλέπουμε πως το μοντέλο δίνει σημαντικά καλύτερα αποτελέσματα τόσο από την βέλτιστη επένδυση buy&hold όσο και από της αφελή, πράγμα που ήταν αναμενόμενο εξαιτίας του καλού ποσοστού σωστής πρόβλεψης που παρατηρήθηκε παραπάνω. Η πρόοδος του κέρδους κάθε φορά φαίνεται στην παρακάτω γραφική ακολουθώντας το μοντέλο πρόβλεψης.



Εικόνα 60. Κέρδος μοντέλου πρόβλεψης για όλες τις τεχνολογίες

Παρατηρείται πως το μοντέλο δεν είναι αλάνθαστο όπως θα περιμέναμε και υπάρχουν περίοδοι που οι προβλέψεις δεν είναι σωστές και οδηγούν σε μείωση του κέρδους, αλλά στην τελική έχουμε κέρδος σε όλες τις περιπτώσεις. Ειδικότερα για το PoW το κέρδος είναι αρκετά μεγαλύτερο γεγονός που προκύπτει από την καλύτερη πρόβλεψη της

αύξησης και μείωσης της μετοχής. Τα περισσότερα κέρδη του μοντέλου προκύπτουν σε περιόδους που υπάρχει συνεχής αύξηση της τιμής των κρυπτονομισμάτων και το μοντέλο προβλέπει σωστά και συνεχίζει να επενδύει κάθε μέρα. Αυτές οι απότομες μεταβολές φαίνονται χαρακτηριστικά στην γραφική αναπαράσταση του κέρδους για το PoW, με τις μεγαλύτερες αυξήσεις τον 01-2023 και τον 03-2023. Πρακτικά τα κέρδη αυτά συμβαίνουν στις αλλαγές επιπέδου της χρονοσειράς και έτσι το μοντέλο καταφέρνει να προβλέπει συνεχώς σωστά την κίνηση της μετοχής. Τα παραπάνω αποτελέσματα είναι αρκετά ενθαρρυντικά και δείχνουν ότι πιθανώς το μοντέλο μπορεί να χρησιμοποιηθεί για καθοδήγηση σε πιθανές χρηματιστηριακές επενδύσεις παρόλης της πρώιμης μορφής του.

6.4 Πρόβλεψη τιμών μεμονωμένων κρυπτονομισμάτων

Το τελευταίο βήμα που μπορεί να μελετηθεί είναι η χρήση του μοντέλου στην πρόβλεψη και επένδυση στα επιμέρους κρυπτονομίσματα με βάση την χρήση του μοντέλου της κάθε τεχνολογίας. Έτσι για παράδειγμα, για την πρόβλεψη του Bitcoin, θα χρησιμοποιηθεί το μοντέλο του αλγορίθμου PoW. Η χρήση του μοντέλου για την πρόβλεψη των κρυπτονομισμάτων και ο υπολογισμός του μέγιστου κέρδους buy&hold, ακολουθούν την ίδια συμπεριφορά που αναφέρθηκε στην προηγούμενη ενότητα και δίνουν τα εξής αποτελέσματα.

Πρωτόκολλο	Κέρδος Μοντέλου	Κέρδος αφελούς buy & hold	Κέρδος βέλτιστου buy & hold
BTC	5.10708	0.592339	7.3888
XMR	7.25022	1.02606	10.9203
DOGE	42.051	0.473324	76.8109
LTC	2.78263	0.830208	18.934
DASH - PoW	6.1313	0.685728	6.32617
ADA	10.1969	0.354762	14.2682
ETH	10.588	0.787945	7.66428
DASH - PoS	3.61209	0.685728	6.32617
TRX	10.546	0.536442	8.29814
XTZ	16.1636	0.93752	33.8455
XLM	25.4137	0.333846	44.8607
XRP	16.0865	0.453101	46.1302
BNB	11.1001	0.699287	13.7705
VET	14.3567	0.476575	11.925

Πίνακας 14. Σύγκριση αποτελεσμάτων επένδυσης σε επιμέρους κρυπτονομίσματα

Η αρχική παρατήρηση είναι ότι για όλα τα κρυπτονομίσματα η αφελής πρόβλεψη δίνει σχεδόν για όλα τα κρυπτονομίσματα αρνητικά αποτελέσματα, και οπότε δεν μπορεί να χρησιμοποιηθεί για προβλέψεις. Παρόλα αυτά είναι το δείγμα της πιο απλής επένδυσης

που μπορεί να γίνει και δίδεται ως παράδειγμα για όλα τα κρυπτονομίσματα. Παρατηρούμε ότι τα αποτελέσματα για τα επιμέρους κρυπτονομίσματα είναι ακόμα καλύτερα από τα αποτελέσματα του κέρδους στους γενικούς δείκτες αλγορίθμων ομοφωνίας. Το μικρότερο κέρδος είναι 2.7 φορές το αρχικό ποσό επένδυσης για το Litecoin, ενώ το μεγαλύτερο είναι 42 φορές το αρχικό ποσό. Τα αποτελέσματα αυτά δείχνουν ότι το μοντέλο κάνει αρκετά καλές προβλέψεις οι οποίες παρόλα αυτά σίγουρα μπορούν να βελτιωθούν. Συγκρίνοντας το κέρδος του buy&hold, εδώ παρατηρούμε μία διαφορά σε σχέση με τους γενικούς δείκτες. Μόνο τρία από τα δεκατέσσερα κρυπτονομίσματα ξεπερνάνε το μέγιστο κέρδος του buy&hold, που σημαίνει ότι οι προβλέψεις για τα επιμέρους κρυπτονομίσματα είναι χειρότερες, πιθανώς γιατί δεν είναι τόσο στοχευμένες. Μία άλλη παρατήρηση είναι ότι η πρόβλεψη με βάση τους δείκτες τείνει να δώσει τυχαία αποτελέσματα. Ενώ θα περιμέναμε τα κρυπτονομίσματα που χρησιμοποιούν αλγορίθμους PoW να έχουν μεγαλύτερο κέρδος ως αναλογία των παραπάνω δεδομένων για τους γενικούς δείκτες, αυτό δεν ισχύει και έχουμε περιπτώσεις που το κέρδος είναι μεγαλύτερο σε κρυπτονομίσματα άλλου είδους. Παρόλα αυτά, η μόνη σύγκριση που μπορεί να θεωρηθεί ακριβής είναι για το DASH, ένα hybrid κρυπτονομίσμα, για το οποίο προβλέπουμε την συμπεριφορά του, τόσο με τον δείκτη PoW όσο και με τον δείκτη PoS. Όπως αναμένουμε ο δείκτης PoW έχει καλύτερα αποτελέσματα παρουσιάζοντας σχεδόν το διπλάσιο κέρδος, και είναι η μόνη σύγκριση που βγάζει συμπεράσματα με νόημα. Πιθανώς να πρέπει να επεκταθεί σε αυτό το κόμματι η δημιουργία γενικών δεικτών με περισσότερα hybrid κρυπτονομίσματα για την εγγενή σχέση των τεχνολογιών στην πρόβλεψη της τιμής τους.

Κεφάλαιο 7. Επίλογος

Ολοκληρώνοντας την μελέτη, αξίζει να γίνει αναφορά στα αποτελέσματα που προκύπτουν όσον αφορά τις προβλέψεις των επιλεγμένων κρυπτονομισμάτων. Βασικός στόχος της εργασίας αποτελούσε η άντληση συμπερασμάτων για τις σχέσεις που υπάρχουν μεταξύ των τεχνολογιών των κρυπτονομισμάτων και η πρόβλεψη των τιμών με την χρήση αυτών. Στην ενότητα αυτή θα περιγράψουν τα συμπεράσματα που προκύπτουν των σχέσεων που προέκυψαν, καθώς και της επίδοσης του μοντέλου πρόβλεψης βάσει των τεχνολογιών. Τέλος θα παρατεθούν κάποιες μελλοντικές προεκτάσεις με σκοπό την καλύτερη άντληση συμπερασμάτων και την βελτίωση του μοντέλου πρόβλεψης κρυπτονομισμάτων.

7.1 Σύνοψη και συμπεράσματα

Βασικός στόχος της εργασίας είναι η ανάλυση και λήψη συμπερασμάτων των παραγόντων που επηρεάζουν την συμπεριφορά της τιμής των κρυπτονομισμάτων. Η ανάλυση αυτή γίνεται διαχωρίζοντας τα κρυπτονομίσματα σε βασικές τεχνολογίες βάσει των αλγορίθμων ομοφωνίας που αυτά χρησιμοποιούν και τελικά συγκρίνεται η συσχέτιση των τεχνολογιών αυτών. Παρόλα αυτά, ένα ακόμα ζητούμενο ήταν και η παραγωγή καλών προβλέψεων της τελικής τιμής των κρυπτονομισμάτων. Συμπεράσματα για την συσχέτιση των τεχνολογιών αντλούνται βάσει της ποιότητας των αποτελεσμάτων πρόβλεψης που προκύπτουν από το μοντέλο σε σχέση με τα δεδομένα που χρησιμοποιούνται κάθε φορά ως είσοδο. Έτσι το μοντέλο κάθε φορά τροφοδοτούνταν με έναν διαφορετικό σύνολο δεδομένων, προσπαθώντας να βρεθεί ο καλύτερος συνδυασμός αυτών. Τα αποτελέσματα που προέκυψαν, δίνουν κάποια συμπεράσματα για την σχέση των τεχνολογιών ως προς την πρόβλεψη των τελικών τιμών των κρυπτονομισμάτων και είναι τα εξής.

- Η καλύτερη γραμμική συσχέτιση διαστήματος ολίσθησης των κρυπτονομισμάτων είναι σε διάστημα ημερών και όχι ωρών ή μηνών. Με ημερήσιο διάστημα ολίσθησης, η χειρότερη γραμμική συσχέτιση είναι της τάξης του 0.7, ενώ με διάστημα ωρών η καλύτερη γραμμική συσχέτιση είναι της τάξης του 0.4, φανερώνοντας της αστάθεια που υπάρχει στα ωριαία δεδομένα και την δυσκολία πρόβλεψης σε αυτό το χρονικό παράθυρο. Παρόλα αυτά για τα ημερήσια δεδομένα η γραμμική συσχέτιση φτάνει σχεδόν το 0.95, δείχνοντας μεγάλη συσχέτιση στην γενική τάση που ακολουθούν τα κρυπτονομίσματα.
- Η ομαδοποίηση κρυπτονομισμάτων φαίνεται να δίνει καλύτερα αποτελέσματα και να εξαλείφει ένα κομμάτι της αστάθειας των κρυπτονομισμάτων. Η ομαδοποίηση που έγινε στην εργασία παράγει αρκετά ικανοποιητικά αποτελέσματα πρόβλεψης της τάσης των κρυπτονομισμάτων σε επίπεδο ημέρας, φτάνοντας στην καλύτερη περίπτωση 77.5% σωστές προβλέψεις. Παρόλα αυτά και στην χειρότερη περίπτωση οι προβλέψεις είναι μεγαλύτερες του 50%, δίνοντας αποτελέσματα καλύτερα από τυχαία.
- Τα κρυπτονομίσματα τύπου PoW έχουν σημαντικά καλύτερο βαθμό πρόβλεψης από τα υπόλοιπα, πιθανώς γιατί είναι αυτά που επηρεάζουν όλα τα υπόλοιπα κρυπτονομίσματα λόγω της δημοτικότητας τους. Το γενικό συμπέρασμα είναι

ότι οι διαφορετικές τεχνολογίες έχουν διαφορετικές συμπεριφορές και διαφορετική γραμμική συσχέτιση και επιρροή μεταξύ τους.

- Τροφοδοτώντας το μοντέλο με δεδομένα ολίσθησης, που παράγουν την καλύτερη γραμμική συσχέτιση, λαμβάνονται ικανοποιητικά αποτελέσματα. Σε συνέχεια αυτού, η προσθήκη ολίσθησης της προηγούμενης ημέρας δίνει ακόμα καλύτερα αποτελέσματα, πιθανώς γιατί αντικατοπτρίζει την γενική τάση της αγοράς. Η προσθήκη παραπάνω δεδομένων χειροτερεύει το μοντέλο προσθέτοντας σε αυτό σφάλματα. Η μεγαλύτερη βελτίωση γίνεται με την επανεκπαίδευση του μοντέλου προσθέτοντας τα πιο πρόσφατα δεδομένα. Τα αποτελέσματα βελτιώνονται επειδή τα πιο πρόσφατα δεδομένα αλλάζουν το επίπεδο της χρονοσειράς πλησιάζοντας περισσότερο στα πραγματικά δεδομένα. Η σημασία της προσθήκης των πρόσφατων δεδομένων φαίνεται από το ποσοστό βελτίωσης των αποτελεσμάτων όπου στην καλύτερη περίπτωση έχουμε μείωση του δείκτη MAPE από 0.179 σε 0.129 (μείωση της τάξης του 5%) και ακόμα και στην χειρότερη περίπτωση έχουμε μείωση του MAPE από 0.159 σε 0.131 (μείωση της τάξης του 2.5%). Η αστάθεια των κρυπτονομισμάτων μετακινεί συνεχώς το επίπεδο της χρονοσειράς, και κάνει την προσθήκη των πιο πρόσφατων δεδομένων αναγκαστική.
- Η προσθήκη γενικών χρηματιστηριακών δεικτών όπως ο S&P500 και ο χρυσός σε πολλές περιπτώσεις δίνει καλύτερα αποτελέσματα ακρίβειας στον δείκτη MAPE, εξομαλύνοντας την συμπεριφορά των προβλέψεων. Παρόλα αυτά, τα ποσοστά σωστής πρόβλεψης είναι χειρότερα στις 7 από τις 8 προβλέψεις των δεικτών ομοφωνίας, πράγμα που δείχνει ότι οι δύο παραπάνω δείκτες χειροτερεύουν τις προβλέψεις με την προσθήκη στο μοντέλο. Από την άλλη, με την προσθήκη του δείκτη του αργού πετρελαίου έχουμε στις περισσότερες των περιπτώσεων, εξομάλυνση των προβλέψεων, με καλύτερα αποτελέσματα του δείκτη MAPE και καλύτερα αποτελέσματα στον δείκτη σωστής πρόβλεψης τάσης. Έτσι συμπεραίνεται ότι ο δείκτης του αργού πετρελαίου μπορεί να χρησιμοποιηθεί ως διαφοροποιητής για το μοντέλο γιατί προσφέρει τόσο καλύτερη εξομάλυνση, όσο και βελτίωση στα αποτελέσματα πρόβλεψης.
- Μέχρι και στην πιο απλή μορφή του μοντέλου, με την χρήση γραμμικής παλινδρόμησης, η πρόβλεψη των επιμέρους κρυπτονομισμάτων είναι καλύτερη από απλές μορφές επενδύσεων σε κρυπτονομίσματα. Το μοντέλο σε πολλές περιπτώσεις προσφέρει παρόμοια αποτελέσματα με την αγορά και πώληση στο καλύτερο σημείο για μία μετοχή, που είναι η ιδανικότερη περίπτωση κέρδους σε απλές επενδύσεις (αφελές και βέλτιστο buy & hold). Ακόμα και σε περιπτώσεις που τα αποτελέσματα δεν πλησιάζουν την ιδανική περίπτωση, το μοντέλο επιφέρει πάντα κέρδος. Στην χειρότερη του επίδοση το μοντέλο επιφέρει κέρδη της τάξης του 2.7 φορές της αρχικής επένδυσης για τα επιμέρους κρυπτονομίσματα.
- Η σημαντικότερη παρατήρηση είναι πως με βάση τα αποτελέσματα της αγοράς στην αρχή και τέλος του συνόλου αξιολόγησης, όλα τα κρυπτονομίσματα εκτός του Monero έχουν πτωτική τάση. Παρόλα αυτά το μοντέλο επιτυγχάνει κέρδος σε όλες τις περιπτώσεις παρόλο που αν επένδυε κάποιος καθ' όλη την διάρκεια του συνόλου αξιολόγησης θα έχανε λεφτά. Αυτό δείχνει ότι οι προβλέψεις του

μοντέλου είναι αρκετά ικανοποιητικές ακόμα και σε αυτό το πρώιμο τους στάδιο και δεν εξαρτώνται από την ανοδική ή καθοδική πορεία των μετοχών.

7.2 Μελλοντικές προεκτάσεις

Η παραπάνω μελέτη αποτελεί αυτοτελές έργο, όπου έγινε εστίαση κυρίως στην ανάδειξη της σχέσης που υπάρχει βάσει των τεχνολογιών των κρυπτονομισμάτων και στην χρήση που μπορεί να γίνει για την πρόβλεψη των τιμών αυτών. Παρόλα αυτά, τα αποτελέσματα μπορούν να βελτιωθούν ώστε να δοθεί μία πληρέστερη εικόνα της χρήσης των αλγορίθμων ομοφωνίας για πρόβλεψη κρυπτονομισμάτων. Η συγκεκριμένη ενότητα μελετά κάποιες πιθανές προεκτάσεις που μπορούν να γίνουν για την παραγωγή μίας ολοκληρωμένης εικόνας.

7.2.1 Έλεγχος ακρίβειας με χρήση μη γραμμικών μοντέλων

Από τις πιο σημαντικές μελλοντικές προεκτάσεις είναι ο έλεγχος ακρίβειας με την χρήση πιο σύνθετων μοντέλων. Σε αρχικό στάδιο μπορούν να ελεγχθούν τα αποτελέσματα με την χρήση μοντέλων χρονοσειρών όπως την εκθετικής εξομάλυνσης και αυτοπαλινδρομικών μοντέλων ARIMA. Σε επόμενο στάδιο, και σύμφωνα με την βιβλιογραφία μπορούν να χρησιμοποιηθούν και μοντέλα μηχανικής μάθησης και νευρωνικών δικτύων. Πιο συγκεκριμένα, είναι πιθανό ότι μοντέλα νευρωνικών δικτύων τύπου RNN και LSTM θα δώσουν καλύτερα αποτελέσματα, καθώς αυτό παρατηρήθηκε στην πρόβλεψη του Bitcoin με την χρήση αυτών.

Γενικά, παρατηρήθηκε ότι η χρήση ενός γραμμικού μοντέλου όπως αυτό της γραμμικής παλινδρόμησης περιορίζει σημαντικά την ακρίβεια σφάλματος και είναι δυσκολότερο να βρεθεί προσεγγιστικά η αναμενόμενη τιμή. Συνεπώς είναι σημαντικό να μελετηθεί και η ακρίβεια με την χρήση μη γραμμικών μοντέλων όπως αυτών που αναφέρθηκαν παραπάνω.

7.2.2 Έλεγχος ακρίβειας σε συναλλαγές πραγματικού χρόνου (real time trading)

Η επίδοση του μοντέλου αξιολογήθηκε ως έναν βαθμό με βάση το κέρδος σε σύγκριση με τα δεδομένα του διαστήματος αξιολόγησης. Παρόλα αυτά οι τιμές αφορούσαν τιμές κλεισίματος και θεωρήθηκε πως δεν υπάρχει ποσοστό προμήθειας σε κάθε συναλλαγή που εκτελούνταν. Η επέκταση του μοντέλου για την χρήση στο χρηματιστήριο με προσομοίωση των συναλλαγών σε πραγματικό χρόνο και τον υπολογισμό της προμήθειας, θα δώσει καλύτερα αποτελέσματα για την δυνατότητα χρήσης του μοντέλου στον πραγματικό κόσμο. Τα μεγαλύτερα συναλλακτήρια επιτρέπουν την άντληση και χρήση των δεδομένων σε πραγματικό χρόνο μέσω API. Έτσι η ακρίβεια του μοντέλου θα αξιολογηθεί με την δημιουργία ενός προγράμματος που συλλέγει δεδομένα και προσομοιώνει την συναλλαγή, σαν να γίνεται μέσω του συναλλακτηρίου.

7.2.3 Δημιουργία πληροφοριακού συστήματος που εκτελεί συναλλαγές με χρήση του μοντέλου

Στα πλαίσια χρήσης του μοντέλου που παρουσιάστηκε, προτείνεται η δημιουργία ενός εργαλείου λήψης αποφάσεων σε πραγματικό χρόνο για την αγορά μετοχών κρυπτονομισμάτων. Η δημιουργία ενός τέτοιου εργαλείου θα μειώσει τα σφάλματα από τον χρήση και θα αυτοματοποιήσει την διαδικασία αγοράς των μετοχών. Παράλληλα, στα πλαίσια του χρηματιστηρίου, όπου οι τιμές των μετοχών αλλάζουν κάθε λεπτό, ένα

αυτοματοποιημένο σύστημα θα αυξήσει την παραγωγή κέρδους με την δυνατότητα άμεσης αγοράς, χωρίς την παρέμβαση του τελικού χρήστη. Πιο συγκεκριμένα, το σύστημα θα δέχεται τα αποτελέσματα του μοντέλου και θα παίρνει αποφάσεις για την επένδυση ή όχι σε μετοχές. Το πρόγραμμα θα εκτελεί μόνο του τη συναλλαγή μέσω ενός συναλλακτηρίου.

7.2.4 Διεύρυνση και μελέτη βέλτιστων ολισθήσεων

Στην παρούσα εργασία έγινε ανάλυση της χρήσης ημερήσιων και ωριαίων ολισθήσεων για την πρόβλεψη και προτιμήθηκε η χρήση ημερήσιων ολισθήσεων. Παράλληλα μελετήθηκαν κάποια αποτελέσματα με την χρήση ημερήσιας ολίσθησης στα πλαίσια μίας και πολλών ημερών.

Μία ενδιαφέρουσα επέκταση του μοντέλου θα ήταν η ενδελεχής ανάλυση του καλύτερου γραμμικού συνδυασμού ημερήσιων ή και ωριαίων ολισθήσεων που δίνουν τα καλύτερα αποτελέσματα πρόβλεψης και του τρόπου εύρεσης αυτών. Η χρήση περισσότερων δεδομένων εισόδου μπορεί να βοηθήσει στην εξομάλυνση των αποτελεσμάτων και η εισαγωγή ωριαίων δεδομένων μαζί με τα ημερήσια μπορεί να δώσει μεγαλύτερη ακρίβεια πρόβλεψης της επόμενης ημερήσιας τιμής μειώνοντας σημαντικά τα σφάλματα.

Από την άλλη, δεν μελετήθηκαν διαστήματα ολίσθησης μεγαλύτερα των δύο μηνών. Ενδεικτικά μπορούν να μελετηθούν διαστήματα τριών, έξι μηνών ή και του ενός χρόνου για την εύρεση, αν αυτή υπάρχει, κυκλικότητας στις χρονοσειρές.

Βιβλιογραφία

- Anon., 2023. *Blockchain.com*. [Ηλεκτρονικό].
- Balcilar, M., Bouri, E., Gupta, R. & Roubaud, D., 2017. Can volume predict Bitcoin returns and volatility? A quantiles-based approach. *Economic Modelling*, Τόμος 64, pp. 74-81.
- Baur, D., Hong, K. & Adrian, L., 2018. Bitcoin: Medium of exchange or speculative assets?. *Journal of International Financial*, pp. 177-189.
- Bernstein, D. J., 2008. Protecting communications against forgery. *Algorithmic Number Theory*.
- Bhaskar, N. D. & Chuen, D. L. K., 2015. Bitcoin Mining Technology. *Handbook of Digital Currency*, pp. 45-66.
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C. & Ljung, G. M., 2015. Time Series Analysis: Forecasting and Control. *John Wiley & Sons*.
- Brown, R. G., 1956. Exponential Smoothing for Predicting Demand. *Little*.
- Chu, J., Chan, S., Nadarajah, S. & Osterrieder, J., 2017. GARCH Modelling of Cryptocurrencies. *Journal of Risk and Financial Management*, 10(4), p. 17.
- Ciaian, P., Rajcaniova, M. & Kancs, d., 2018. Virtual relationships: Short-and long-run evidence from BitCoin and altcoin markets. *Journal of International Financial Markets, Institutions and Money*, Τόμος 52, pp. 173-195.
- Clark, A., 2022. Ethereum's Energy Revamp Is No Guarantee of Global Climate Gains. *Bloomberg*.
- Corbet, S., Lucey, B., Urquhart, A. & Yarovaya, L., 2019. Cryptocurrencies as a financial asset: A systematic analysis. *International Review of Financial Analysis*, pp. 182-199.
- Corbet, S. και συν., 2018. Exploring the dynamic relationships between cryptocurrencies and other financial assets. *Economics Letters*, Τόμος 165, pp. 28-34.
- Coventry, A., 2012. NooShare: A decentralized ledger of shared computational resources. *Self Published*.
- Crosby, M., Pattanayak, P., Verma, S. & Kalyanaraman, V., 2016. Blockchain technology: Beyond bitcoin. *Applied Innovation*, 2(6-10), p. 71.
- Dzulfikar, F. & Susanto, A., 2020. Implementation of Smart Contracts Ethereum Blockchain in Web-Based Electronic Voting. *Jurnal Transformatika*.
- ethos.dev, 2022. The Beacon Chain Ethereum 2.0 explainer you need to read first. *ethos.dev*.
- European Central Bank, 2018. Virtual Currency Schemes - a further analysis.
- Gardner, E. J., 2015. Conservative forecasting with the damped trend. *Journal of Business Research*.

- Gardner, E. J. & McKenzie, E., 1985. Forecasting trends in time series. *Management science*.
- Gkillas, K., Bekiros, S. & Siriopoulos, C., 2018. Extreme correlation in cryptocurrency markets.
- Granger, C. W. J. & Newbold, P., 2014. Forecasting economic time series. *Academic Press*.
- Holt, C. C., 1957. Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, 20(1), pp. 5-10.
- Jakobsson, M. & Juels, A., 1999. Proofs of work and bread pudding protocols. Στο: *Secure information networks*. Boston, MA: Springer, pp. 258-272.
- Kumah, S. P. & Jones, O.-M., 2022. Do cryptocurrencies and crude oil influence each other? Evidence from wavelet-based quantile-in-quantile approach. *Cogent Economics & Finance* , 10(1).
- Lawuobahsumo, K. K., Algieri , B., Iania, L. & Leccadito, A., 2022. Exploring Dependence Relationships between Bitcoin and Commodity Returns: An Assessment Using the Gerber Cross-Correlation. *Commodities*, 1(1), pp. 34-49.
- Maiti, M., 2022. Dynamics of bitcoin prices and energy consumption. *Chaos, Solitons & Fractals: X*, Τόμος 9.
- McGleenon, B., 2022. Why is this crypto token rising amid crash and FTX collapse?. *Yahoo Finance*.
- Menezes, A., van Oorschot, P. & Vanstone, S., 1996. *Handbook of applied cryptography*. s.l.:CRC Press.
- Nakamoto, S., 2008. Bitcoin: A Peer-to-Peer Electronic Cash System. *Decentralized Business Review*, Τόμος 21260.
- Osterrieder, J. & Lorenz, J., 2017. A statistical risk assessment of Bitcoin and its extreme tail behavior. *Annals of Financial Economics*, 12(01).
- Sloughter, J. M. L., Raftery, A. E., Gneiting, T. & Fraley, C., 2007. Probabilistic Quantitative Precipitation Forecasting Using Bayesian Model Averaging. *Monthly Weather Review*.
- Söylemez, Y. & Gürsoy, S., 2022. An analysis of the causality relationship between bitcoin electricity consumption, price and volume. *Journal of Research in Business*, 7(1).
- Tasca, P. & Tessone, C., 2017. Taxonomy of Blockchain Technologies:Principles of Identification and Classification. *Principles of identification and classification*.
- The Economist, 2015. Blockchains: The great chain of being sure about things. *The Economist*.
- Urquhart, A. & Zhang, H., 2019. Is Bitcoin a hedge or a safe haven for currencies? An intra day analysis. *International Review of Financial Analysis*, Τόμος 63, pp. 49-57.