



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ ΔΙΑΤΑΞΕΩΝ

ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ

ΩΡΙΑΙΑ ΠΡΟΒΛΕΨΗ ΤΙΜΗΣ ΚΛΕΙΣΙΜΑΤΟΣ ΚΡΥΠΤΟΝΟΜΙΣΜΑΤΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΕΛΕΥΘΕΡΙΟΣ ΣΤ. ΚΑΛΟΓΙΑΝΝΗΣ

Επιβλέπων : Βασίλειος Ασημακόπουλος

Καθηγητής ΕΜΠ

Υπεύθυνος: Ευάγγελος Σπηλιώτης

Διδάκτωρ ΕΜΠ

Αθήνα, Ιανουάριος 2020



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ ΔΙΑΤΑΞΕΩΝ
ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ

ΩΡΙΑΙΑ ΠΡΟΒΛΕΨΗ ΤΙΜΗΣ ΚΛΕΙΣΙΜΑΤΟΣ
ΚΡΥΠΤΟΝΟΜΙΣΜΑΤΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΕΛΕΥΘΕΡΙΟΣ ΣΤ. ΚΑΛΟΓΙΑΝΝΗΣ

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 24^η Ιανουαρίου 2020

.....
Β. Ασημακόπουλος,
Καθηγητής ΕΜΠ

.....
Ι. Ψαρράς,
Καθηγητής ΕΜΠ

.....
Δ. Ασκούνης
Καθηγητής ΕΜΠ

Αθήνα, Ιανουάριος 2020

.....

© ΕΛΕΥΘΕΡΙΟΣ ΣΤ. ΚΑΛΟΓΙΑΝΝΗΣ

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών ΕΜΠ.

Copyright © ΕΛΕΥΘΕΡΙΟΣ ΣΤ. ΚΑΛΟΓΙΑΝΝΗΣ, 2020

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 αφορά στην ανάλυση δεδομένων κρυπτονομισμάτων με σκοπό την εξαγωγή προβλέψεων και την αξιοποίησή τους για την εκτέλεση επικερδών συναλλαγών. Στόχος της μελέτης αυτής, είναι η ανάπτυξη μεθοδολογίας για τις πιο αποδοτικές προβλέψεις της τιμής κλεισίματος του κρυπτονομίσματος bitcoin σε ωριαία βάση και την αξιολόγησή της ως επενδυτικό εργαλείο.

Αρχικά, παρουσιάζεται η τεχνολογική καινοτομία του blockchain και της εφαρμογής του στα κρυπτονομίσματα. Συγκεκριμένα, αναλύονται τα βασικά χαρακτηριστικά του blockchain καθώς και οι τομείς που η υλοποίησή του βρίσκει εύπορο έδαφος για μελλοντική εξέλιξη. Στη συνέχεια, παρουσιάζονται εκτενέστερα τα χαρακτηριστικά των δύο κύριων κρυπτονομισμάτων που κατέχουν το μεγαλύτερο μερίδιο αγοράς σήμερα, bitcoin και ethereum, καθώς και η ευρέως εφαρμοσμένη μέθοδος στο χρηματοοικονομικό κλάδο, αυτή των αλγοριθμικών συναλλαγών.

Σε επόμενο κεφάλαιο, παρουσιάζεται συνοπτικά το γνωστικό υπόβαθρο και δίνονται οι ορισμοί βασικών εννοιών που χρησιμοποιούνται στην παρούσα εργασία, τόσο για την καλύτερη κατανόηση, όσο και για την περαιτέρω αξιοποίησή της, και τέλος γίνεται μία εισαγωγή στις μεθόδους προβλέψεων.

Αναλυτικότερα, περιγράφονται οι μαθηματικές σχέσεις και η ερμηνεία των μεθόδων που αξιολογήθηκαν στο πλαίσιο της παρούσας εργασίας, οι οποίες είναι οι εξής: Απλοϊκή μέθοδος, Μέθοδος Κινητών Μέσων Όρων, Μέθοδος Εκθετικής Εξομάλυνσης και Αυτοπαλινδρομικά Μοντέλα Κινητού Μέσου Όρου. Επίσης, αποτυπώνεται η στρατηγική της χρονικής συνάθροισης, καθώς και ο τρόπος με τον οποίο η αξιοποίησή της επιδρά θετικά στα αποτελέσματα των προβλέψεων.

Στη συνέχεια, περιγράφεται η εκτέλεση της πειραματικής διαδικασίας που ακολουθήθηκε καθ' όλη τη διάρκεια της μελέτης και συγκεκριμένα, από τη συλλογή των δεδομένων μέχρι τη χρήση των προβλέψεων για τη λήψη αποφάσεων πραγματοποίησης ή μη συναλλαγών στο δείγμα. Η πειραματική διαδικασία υλοποιήθηκε με χρήση της γλώσσας προγραμματισμού R (statistics), με την οποία αναλύθηκαν ιστορικά δεδομένα του bitcoin με τη μέθοδο της χρονικής συνάθροισης και παρήχθησαν ωριαίες προβλέψεις για το επιλεχθέν δείγμα. Οι τελευταίες αξιολογήθηκαν βάσει των σφαλμάτων τους και τελικά επιλέχθηκε η μέθοδος με τα καλύτερα αποτελέσματα, η απλή εκθετική εξομάλυνση, με χρήση βέλτιστων συντελεστών στα διάφορα επίπεδα χρονικής συνάθροισης.

Σε συνέχεια των παραπάνω, παρουσιάζονται τα αποτελέσματα των πειραμάτων εκτέλεσης που τεκμηριώνουν τις αποφάσεις που πάρθηκαν κατά την πειραματική διαδικασία, αλλά και τα αποτελέσματα της αξιοποίησης της μεθοδολογίας ως ένα στατιστικό επενδυτικό εργαλείο.

Τέλος, συνοψίζονται τα συμπεράσματα της παρούσας διπλωματικής εργασίας και επισημαίνονται τα σημεία που κρίνονται κρίσιμα για την επιτυχή χρήση της μεθοδολογίας που αναπτύχθηκε. Ολοκληρώνοντας την παρούσα μελέτη, προτείνονται μελλοντικές προεκτάσεις ώστε να καταστεί το μοντέλο ένα ολοκληρωμένο επενδυτικό πακέτο για αυτοματοποιημένες επικερδής συναλλαγές κρυπτονομισμάτων.

ABSTRACT

The present thesis deals with the analysis of cryptocurrency data for the purpose of extracting forecasts and exploiting them to execute profitable transactions. The purpose of this study is to develop a methodology that generates accurate forecasts about the closing price of bitcoin on an hourly basis and evaluate this methodology as an investment tool.

First, the technological innovation of blockchain and its application to cryptocurrencies is presented. In particular, the key features of the blockchain are analyzed, as well as the areas where its implementation provides a good basis for future development. Then, the features of the two main coins, bitcoin and etherium, as well as the widely used method in the financial sector, that of algorithmic transactions, are presented in greater detail.

Additionally, the cognitive background required for both understanding and utilizing this study is approached and an introduction to forecasting methods is made. In more detail, the mathematical relationships and interpretation of the evaluated methods are described, which are the Naive, the Moving Averages, the Exponential Smoothing method, and the Autoregression Moving Average Models. It also illustrates the temporal aggregation strategy and shows how its utilization in time series analysis can have a positive effect on forecasting accuracy.

Subsequently, the entire execution of the experimental procedure from data collection to the use of predictive trading decisions throughout the sample, is described. The experiment was carried out using the R (statistics) programming language, in which historical bitcoin data were analyzed by exploiting the temporal aggregation method, and hourly forecasts for the selected sample are provided. The latter were evaluated on the basis of their errors and finally the method with the best results, the simple exponential smoothing, was chosen using optimal coefficients at the different temporal levels. Following the description, all experiment results are presented, including those which document the accuracy of the model, as well as the results of the methodology implementation as statistical investment adviser for the entire sample.

Last but not least, the conclusions of this thesis are summarized, pointing out the key points for the success of using the methodology which was developed. Future extensions are also proposed to make the model an integrated investment package for automated profitable cryptocurrency trading.

Πρόλογος

Η διπλωματική αυτή εργασία εκπονήθηκε στα πλαίσια των ερευνητικών δραστηριοτήτων της Μονάδας Προβλέψεων και Στρατηγικής κατά το ακαδημαϊκό έτος 2019-2020. 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 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. Ιωάννη Ψαρρά 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, αξίζει να σημειωθεί ότι το έναυσμα για την επιλογή του ερευνητικού ζητήματος που πραγματεύεται αυτή η μελέτη πηγάζει από την ενασχόλησή μου με την start-up εταιρεία “Ducimus”. Στο πλαίσιο αυτό, θα ήθελα να ευχαριστήσω από καρδιάς τους συμφοιτητές μου Πιστιόλα Νικόλαο - Στρατή, Γκλαντζλούνη Κωσταντίνο και Μπιρμπάκο Άγγελο για τη σύντομη αλλά εποικοδομητική πορεία που είχαμε σε όλο τον κύκλο ζωής της start-up. Επίσης, θα ήθελα να ευχαριστήσω την οικογένεια μου για τη στήριξη που μου παρείχε σε όλη τη διάρκεια των σπουδών μου.

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

Περίληψη.....	4
ABSTRACT	5
Πρόλογος.....	6
ΠΕΡΙΕΧΟΜΕΝΑ ΕΙΚΟΝΩΝ	9
ΠΕΡΙΕΧΟΜΕΝΑ ΠΙΝΑΚΩΝ	11
1. Εισαγωγή στα κρυπτονομίσματα	13
1.1. Η έννοια του blockchain	13
1.1.1. Hash function.....	13
1.1.2. Block	13
1.1.3. Κόμβοι	13
1.1.4. Blockchain.....	13
1.2. Εφαρμογές του blockchain	15
1.2.1. Ασφάλιση.....	15
1.2.2. Διακυβέρνηση	15
1.2.3. Ψηφοφορία	15
1.2.4. Λιανικό εμπόριο	16
1.2.5. Αγορά ακινήτων (Real Estate)	16
1.3. Βασικότερα κρυπτονομίσματα.....	16
1.3.1. Bitcoin.....	17
1.3.2. Ethereum	24
1.4. Αλγοριθμικές Συναλλαγές.....	27
1.4.1. Στρατηγικές Αλγοριθμικών Συναλλαγών.....	27
1.4.2. Συναλλαγές υψηλής συχνότητας.....	29
2. Εισαγωγή στις προβλέψεις.....	33
2.1. Γενικά για τις προβλέψεις	33
2.2. Χαρακτηριστικά Χρονοσειρών.....	35
2.3. Βασικές Μέθοδοι Πρόβλεψης.....	38
2.3.1. Απλοϊκή Μέθοδος (Naïve).....	39
2.3.2. Μέθοδος Κινητού Μέσου Όρου	39

2.3.3.	Εκθετική εξομάλυνση	40
2.3.4.	Μοντέλα αυτοσυσχέτισης (Arima model)	43
2.4.	Ακρίβεια προβλέψεων (Σφάλματα)	47
3.	Μέθοδοι Συνάθροισης Δεδομένων (Aggregation)	51
3.1.	Ιεραρχική Συνάθροιση	51
3.2.	Χρονική Συνάθροιση	53
3.2.1.	Μη Επικαλυπτόμενη Χρονική Συνάθροιση	53
3.2.2.	Επικαλυπτόμενη Χρονική Συνάθροιση (.....)	55
3.2.3.	Αλγόριθμος προβλέψεων πολλαπλής συνάθροισης.....	56
4.	Πειραματική Διαδικασία	59
4.1.	Στόχος της Πειραματικής Διαδικασίας	59
4.2.	Επιλογή Επενδυτικού Προϊόντος	59
4.3.	Συλλογή και Εξόρυξη Δεδομένων	59
4.4.	Διαχείριση Ασυνήθιστων Τιμών	60
4.5.	Επιλογή Μεθόδου Πρόβλεψης.....	62
4.6.	Χρονική Συνάθροιση της χρονοσειράς (Temporal Aggregation)	62
4.7.	Παραγωγή προβλέψεων σε διαφορετικά επίπεδα χρονικής συνάθροισης 66	
4.8.	Συνδυασμός διαφορετικών επιπέδων χρονικής συνάθροισης	67
4.9.	Χρήση του μοντέλου πρόβλεψης ως εργαλείου αγοροπωλησίας	68
5.	Παρουσίαση Αποτελεσμάτων.....	71
5.1.	Αποτελέσματα επιλογής μεθόδου πρόβλεψης	71
5.1.1.	Μέθοδος Naive	71
5.1.2.	Μέθοδος Holt	72
5.1.3.	Μέθοδος SES	73
5.1.4.	Μέθοδος Moving Average.....	74
5.1.5.	Μέθοδος ARIMA.....	74
5.2.	Αποτελέσματα προβλέψεων χρονικής συνάθροισης.....	75
5.2.1.	Επίπεδο συνάθροισης (aggregation) 1 λεπτού	75
5.2.2.	Επίπεδο συνάθροισης (aggregation) 5 λεπτού	76
5.2.3.	Επίπεδο συνάθροισης (aggregation) 15 λεπτών	78

5.2.4.	Επίπεδο συνάθροισης (aggregation) 30 λεπτών	79
5.2.5.	Επίπεδο συνάθροισης (aggregation) 60 λεπτών	80
5.3.	Αποτελέσματα προβλέψεων με MAPA	82
5.4.	Αποτελέσματα πειράματος εφαρμογής	83
6.	Συμπεράσματα και προοπτικές περαιτέρω έρευνας	87
6.1.	Συμπεράσματα	87
6.2.	Μελλοντικές Προεκτάσεις	88
6.2.1.	Γενικά.....	88
6.2.2.	Δημιουργία εργαλείου που εκτελεί τις συναλλαγές.....	88
6.2.3.	Δημιουργία εργαλείου που συλλέγει τα δεδομένα	88
6.2.4.	Εναλλακτικές μέθοδοι πρόβλεψης.....	88
	Βιβλιογραφία.....	91
	Ηλεκτρονικές Πηγές.....	92
	Παραρτήματα	93
	Παράρτημα κώδικα R	93
	Π1. Βιβλιοθήκες.....	93
	Π2. Εισαγωγή Δεδομένων	93
	Π3. Μέθοδοι Πρόβλεψης	94

ΠΕΡΙΕΧΟΜΕΝΑ ΕΙΚΟΝΩΝ

Εικόνα 1.	Απεικόνιση σχεδίου ιεραρχίας ενός block.....	14
Εικόνα 2.	Μερίδιο αγοράς που κατέχει το κάθε κρυπτονόμισμα.....	16
Εικόνα 3.	Ρυθμός παραγωγής Bitcoin	20
Εικόνα 4.	Σύγκριση κατανάλωσης ενέργειας μεταξύ κρατών (Πηγή: BitcoinEnergyConsumption.com).....	24
Εικόνα 5.	Σύγκριση Bitcoin network με VISA network (Πηγή: BitcoinEnergyConsumption.com).....	24
Εικόνα 6.	HFT σε ποσοστό των ημερήσιων συναλλαγών μετοχών στις Η.Π.Α (Πηγή: Financial Times).....	30
Εικόνα 7.	Κόστος εκτέλεσης εντολής 500.000\$.....	31
Εικόνα 8.	Απεικόνιση χαρακτηριστικού παραδείγματος αποσύνθεσης χρονοσειράς	35

Εικόνα 9. Απεικόνιση της τιμής του κρυπτονομίσματος bitcoin για την περίοδο Αύγουστος – Δεκέμβριο 2018 (Πηγή: Highcharts.com).....	37
Εικόνα 10. Απεικόνιση της μεταβλητότητας (volatility)του κρυπτονομίσματος bitcoin για την περίοδο Αύγουστος – Δεκέμβριο 2018 (Πηγή: Highcharts.com).....	38
Εικόνα 11. Αποτύπωση μη-επικαλυπτόμενης χρονικής συνάθροισης.	54
Εικόνα 12. Απεικόνιση μη-επικαλυπτόμενης χρονικής συνάθροισης.....	55
Εικόνα 13. Απεικόνιση της μεθοδολογίας MAPA.	57
Εικόνα 14. Χαρακτηριστικό παράδειγμα ασυνήθιστων τιμών πριν την αντικατάστασή τους.	61
Εικόνα 15. Χαρακτηριστικό παράδειγμα ασυνήθιστων τιμών μετά την αντικατάστασή τους.	61
Εικόνα 16. Απεικόνιση όλων των δεδομένων που συλλέχθηκαν.....	63
Εικόνα 17. Set 1 min	64
Εικόνα 18. Set 5 mins.....	64
Εικόνα 19. Set 15 mins.....	65
Εικόνα 20. Set 30 mins.....	65
Εικόνα 21. Set 60 mins.....	66
Εικόνα 22. Διάγραμμα Ροής(Flow chart) του αλγορίθμου εφαρμογής του μοντέλου.	69
Εικόνα 23. Γραφική απεικόνιση προβλέψεων (μπλε χρώμα) / πραγματικών τιμών , 50 ωρών για το Set 1 min.....	76
Εικόνα 24. Γραφική απεικόνιση προβλέψεων (κόκκινο χρώμα) / πραγματικών τιμών , 50 ωρών για το Set 5 min.	77
Εικόνα 25. Γραφική απεικόνιση προβλέψεων (ροζ χρώμα) / πραγματικών τιμών , 50 ωρών για το Set 30 min.....	80
Εικόνα 26. Γραφική απεικόνιση προβλέψεων (πράσινο χρώμα) / πραγματικών τιμών , 50 ωρών για το Set 60 min.	81
Εικόνα 27. Γραφική απεικόνιση προβλέψεων (πορτοκαλί χρώμα) / πραγματικών τιμών , 50 ωρών για το συνδυασμό όλων των επιπέδων χρονικής συνάθροισης.....	83
Εικόνα 28. Απεικόνισης των κερδών στην εξέλιξη του χρόνου.....	84

ΠΕΡΙΕΧΟΜΕΝΑ ΠΙΝΑΚΩΝ

Πίνακας 1. Αποτελέσματα έρευνας με τη μέθοδο Naïve	71
Πίνακας 2. Αποτελέσματα έρευνας με τη μέθοδο Holt	72
Πίνακας 3. Αποτελέσματα έρευνας με τη μέθοδο Ses.....	73
Πίνακας 4. Αποτελέσματα έρευνας με τη μέθοδο Moving Average	74
Πίνακας 5. Αποτελέσματα έρευνας με τη μέθοδο ARIMA	74
Πίνακας 6. Αποτελέσματα επιπέδου συνάθροισης 1 λεπτού	75
Πίνακας 7. Μέσα απόλυτα σφάλματα για ολόκληρο το δείγμα (1 λεπτού)	75
Πίνακας 8. Αποτελέσματα επιπέδου συνάθροισης 5 λέπτου	76
Πίνακας 9. Μέσα απόλυτα σφάλματα για ολόκληρο το δείγμα (5 λεπτών).....	77
Πίνακας 10. Αποτελέσματα επιπέδου συνάθροισης 15 λέπτου	78
Πίνακας 11. Μέσα απόλυτα σφάλματα για ολόκληρο το δείγμα (15 λεπτών).....	78
Πίνακας 12. Αποτελέσματα επιπέδου συνάθροισης 30 λέπτου	79
Πίνακας 13. Μέσα απόλυτα σφάλματα για ολόκληρο το δείγμα (30 λεπτών).....	79
Πίνακας 14. Αποτελέσματα επιπέδου συνάθροισης 60 λεπτών.....	80
Πίνακας 15. Μέσα απόλυτα σφάλματα για ολόκληρο το δείγμα (60 λεπτών).....	81
Πίνακας 16. Αποτελέσματα προβλέψεων με MAPA	82
Πίνακας 17. Πίνακας δεδομένων εισαγωγής	83
Πίνακας 18. Πίνακας αποτελεσμάτων.....	84

1. Εισαγωγή στα κρυπτονομίσματα

1.1. Η έννοια του blockchain

Το blockchain είναι μία τεχνολογική εξέλιξη στον κλάδο της επιστήμης των υπολογιστών και της κρυπτογραφίας που παρέχει τη δυνατότητα στους χρήστες του να πιστοποιούν μία μορφή συναλλαγών με απόλυτη αδιαβλητότητα και με απουσία κεντρικού ελέγχου. Παρακάτω παρουσιάζεται η δομή του blockchain, υποθέτοντας μία διαδικασία συναλλαγών οι οποίες καταγράφονται σε ένα αρχείο και πιστοποιούνται μέσω της τεχνολογίας αυτής.

1.1.1. Hash function

Κάθε συναλλαγή που πραγματοποιείται καταγράφεται στο αρχείο μέσω μιας συνάρτησης κρυπτογράφησης (Cryptographic Hash Function) με μοναδική κωδικοποίηση σταθερού μήκους ανεξαρτήτως μεγέθους και προορισμού της συναλλαγής.

1.1.2. Block

Κάθε τέτοιο αρχείο που περιλαμβάνει έναν αριθμό συναλλαγών με μέγιστη χωρητικότητα που εξαρτάται από την αρχιτεκτονική που χρησιμοποιείται αποτελεί ένα block. Ο αριθμός αυτός διαφέρει ανά κρυπτονομίσμα, για παράδειγμα η μέση χωρητικότητα που καταλαμβάνει μία συναλλαγή με bitcoin είναι 570 bytes , δηλαδή περίπου 3500 συναλλαγές ανά block, δεδομένου του 1MB ανώτατου ορίου μνήμης του bitcoin. Ενώ στο ethereum, σε ότι αφορά τα δεδομένα, το μέγεθος του block είναι λιγότερο από 2KB και έχει περίπου 70 συναλλαγές καταγεγραμμένες.

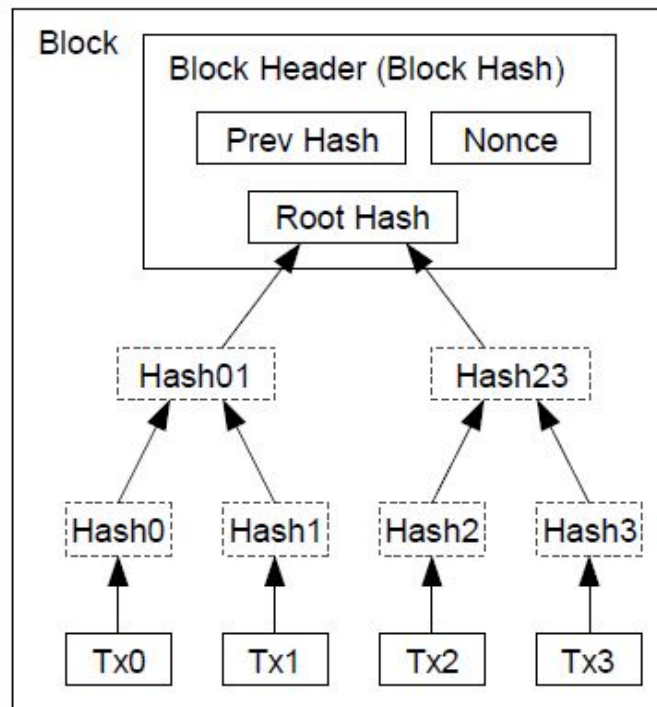
1.1.3. Κόμβοι

Όλα τα blocks υπάρχουν, αναπαράγονται και μεταφέρονται μεταξύ υπολογιστών, που αποτελούν ουσιαστικά τους χρήστες της τεχνολογίας blockchain και ονομάζονται κόμβοι(nodes). Κάθε κόμβος έχει όλο το ιστορικό όλων των blocks που έχουν δημιουργηθεί από την αρχή της δημιουργίας του συγκεκριμένου blockchain. Όλοι οι κόμβοι συμβάλουν στη διαδικασία πιστοποίησης των συναλλαγών και εφόσον η πλειοψηφία συμφωνεί ότι η συναλλαγή είναι έγκυρη , τότε εγγράφεται στο αντίστοιχο block.

1.1.4. Blockchain

Όλοι οι παραπάνω κόμβοι είναι συνδεδεμένοι μεταξύ τους μέσω του διαδικτύου και αποτελούν ουσιαστικά μία αλυσίδα από όλα τα blocks, δηλαδή το blockchain. Το blockchain ανανεώνεται κάθε 10 λεπτά αυτόματα, χωρίς κάποια εντολή ή κεντρικό έλεγχο. Κάθε φορά που γεμίζει ένα block με συναλλαγές και πρέπει να γίνει καινούργια εγγραφή το blockchain παράγει ένα νέο block. Πρόκειται ουσιαστικά για ένα αρχείο καταγραφής ή ψηφιακό ημερολόγιο (digital ledger) στο οποίο όλοι οι χρήστες (nodes)

έχουν όλο το ιστορικών των συναλλαγών καταγεγραμμένο, για αυτό και συχνά αναφέρεται και ως αποκεντροποιημένη (decentralized) οικονομία .



Εικόνα 1. Απεικόνιση σχεδίου ιεραρχίας ενός block

Το blockchain χωρίζεται σε 3 βασικές κατηγορίες.

1.1.4.1. Δημόσιο (Public) Blockchain

Αποτελεί την πιο διαδεδομένη μορφή της τεχνολογίας αφού είναι η βάση πάνω στην οποία λειτουργούν τα κρυπτονομίσματα. Στο πλαίσιο αυτό , οποιοσδήποτε έχει τη δυνατότητα πρόσβασης και άδεια να δημιουργήσει κόμβους , να πραγματοποιήσει συναλλαγές και να δει ολόκληρη των blocks από την δημιουργία του έως τώρα.

1.1.4.2. Ιδιωτικό (Private) Blockchain

Όπως είναι προφανές από την ονομασία του, πρόκειται για ένα κλειστό σύστημα που ένας χρήστης επιλέγει ποιος ή ποιοι μπορούν να έχουν πρόσβαση να προσπελάσουν το Blockchain, να εκτελέσουν συναλλαγές ή να τροποποιήσουν το σύστημα. Αρκετοί υποστηρίζουν ότι δεν θα έπρεπε να ονομάζεται έτσι , δεδομένου ότι αντιτίθεται στην ιδιοσυγκρασία της αποκεντροποιημένης φύσης της τεχνολογίας αυτής. Χαρακτηριστικό παράδειγμα αυτού είναι η πλατφόρμα «MultiChain» (<https://www.multichain.com/>) , η οποία παρέχει τη δυνατότητα σε εταιρείες, κυρίως πολυεθνικές, να φτιάξουν ιδιωτικά blockchains , με σκοπό την βελτίωση των συναλλαγών στο εσωτερικό τους.

1.1.4.3. Consortium ή Federated (Ομοσπονδιακό) Blockchain

Αυτή η μορφή του blockchain αποτελεί κατά μία έννοια συγκερασμό των προηγούμενων δύο. Πιο συγκεκριμένα, αντί να είναι ένας επικεφαλής χρήστης, υπάρχει μία ομάδα χρηστών που παίρνει αποφάσεις και έχει όλα τα δικαιώματα του δικτύου. Για παράδειγμα η μη κερδοσκοπικός οργανισμός Energy Web Foundation (<https://www.energyweb.org/about/what-we-do/>), αποτελεί το μεγαλύτερο οικοσύστημα blockchain στον κλάδο της ενέργειας με σκοπό οι χρήστες του να μοιράζονται έμπιστα τεχνολογικά επιτεύγματα, πληροφορίες και υποστήριξη.

1.2. Εφαρμογές του blockchain

Η τεχνολογία του blockchain εμφανίστηκε μέσω της πρώτης και πιο διαδεδομένης εφαρμογής τους, τα κρυπτονομίσματα. Παρόλα αυτά η τεχνολογία αυτή μπορεί να βρει εφαρμογή σε αρκετά διαφορετικούς τομείς. Παρακάτω παρουσιάζονται συνοπτικά ορισμένες εφαρμογές του blockchain.

1.2.1. Ασφάλιση

Η παγκόσμια ασφαλιστική αγορά βασίζεται στη διαχείριση εμπιστοσύνης, οπότε το blockchain μπορεί να χρησιμοποιηθεί για την επαλήθευση πολλών τύπων δεδομένων σε ασφαλιστικά συμβόλαια, όπως η ταυτότητα του ασφαλισμένου. Η Synarhea (<https://synarhea.com>), η πρώτη ελληνική εφαρμογή blockchain «κατασκευάζει» smart contracts για την ασφαλιστική αγορά.

1.2.2. Διακυβέρνηση

Η εφαρμογή συστημάτων βασισμένων σε blockchain μπορεί να μειώσει σημαντικά τη γραφειοκρατία και να αυξήσει την ασφάλεια, την αποτελεσματικότητα και τη διαφάνεια των κυβερνητικών διαδικασιών. Για παράδειγμα, το Ντουμπάι έχει θέσει στόχο να εγγράψει όλα τα κυβερνητικά του έγγραφα μέσω του blockchain μέχρι το 2020 (<https://www.smartdubai.ae/initiatives/blockchain>). Η Stampdio (<https://stampd.io>), η πρώτη ελληνική εφαρμογή blockchain που δημιουργεί κρυπτογραφικό αποτύπωμα ενός εγγράφου στο blockchain ώστε να εξασφαλίζεται η μοναδικότητα και να ελαχιστοποιούνται οι πιθανότητες να πλαστογραφηθεί από άλλον (ένα παράδειγμα εφαρμογής του blockchain στον κλάδο των συμβολαιογράφων).

1.2.3. Ψηφοφορία

Η τεχνολογία blockchain μπορεί να χρησιμοποιηθεί για την εγγραφή ψηφοφόρων, την εξακρίβωση ταυτότητας και την ηλεκτρονική καταμέτρηση των ψήφων για να διασφαλιστεί ότι υπολογίζονται μόνο οι νόμιμες ψήφοι. Τα Democracy Earth (<https://democracy.earth/>) και Follow My Vote (<https://followmyvote.com/>) αποσκοπούν στην δημιουργία πιο δίκαιων και δημοκρατικών ηλεκτρονικών συστημάτων ψηφοφορίας βασισμένα σε blockchain.

1.2.4. Λιανικό εμπόριο

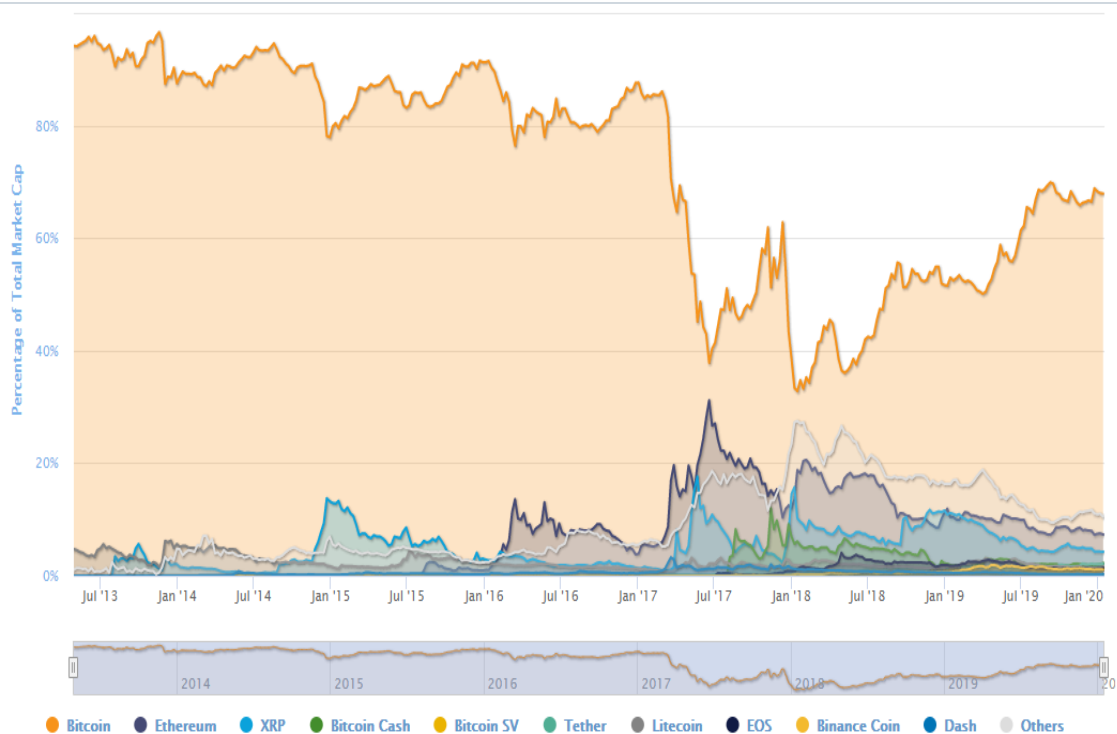
Οι αποκεντρωμένες υπηρεσίες λιανικής που βασίζονται σε blockchain συνδέουν αγοραστές και πωλητές χωρίς μεσάζοντα και συναφή τέλη. Σε αυτές τις περιπτώσεις, η εμπιστοσύνη προέρχεται από τα «έξυπνα συμβόλαια»(smart contracts), την ασφάλεια των ανταλλαγών και τα ενσωματωμένα συστήματα διαχείρισης διαδικτυακής φήμης. Το OpenBazaar (<https://openbazaar.org/>) είναι μια startup που δραστηριοποιείται σε αυτόν τον τομέα.

1.2.5. Αγορά ακινήτων (Real Estate)

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. Η Ubitquity (<https://www.ubitquity.io/>) είναι μια πλατφόρμα για την τήρηση αρχείων ακινήτων στο blockchain.

1.3. Βασικότερα κρυπτονομίσματα

Είναι ευρέως γνωστό ότι ο κύριος λόγος διάδοσης και εξέλιξης της τεχνολογίας blockchain είναι η εφαρμογή του στα κρυπτονομίσματα. Παρακάτω παρατίθεται μία απεικόνιση σε ένα διάγραμμα του μεριδίου αγοράς που κατέχει κάθε κρυπτονομίσμα για τη χρονική περίοδο Ιανουάριος 2013 έως Ιανουάριος 2020.



Εικόνα 2. Μεριδίο αγοράς που κατέχει το κάθε κρυπτονομίσμα

Όπως είναι εμφανές το bitcoin κυριαρχεί ολοκληρωτικά σε όλη την περίοδο και ακολουθεί με μεγάλη απόσταση το ethereum. Δεδομένου αυτού, αλλά και ότι το ethereum εισήγαγε μία άλλη διάσταση στο κόσμο του blockchain με την υλοποίηση των «έξυπνων συμβολαίων» (smart contracts), τα δύο αυτά κρυπτονομίσματα είναι αυτά που θα αναλυθούν εκτενέστερα παρακάτω.

1.3.1. Bitcoin

Το μπιτκοϊν (Bitcoin) είναι ένα ψηφιακό νόμισμα, χωρίς υλική υπόσταση, που δημιουργήθηκε στη βάση της τεχνολογίας του blockchain και εισήγαγε μία νέα εποχή τόσο τομέα της νομισματικής επιστήμης και του χρηματιστηρίου, όσο και στον κλάδο της πληροφορικής. Αποτελεί ουσιαστικά την πρώτη εφαρμογή της τεχνολογίας blockchain και τον βασικό πυρήνα μιας έννοιας που ονομάζεται «κρυπτονόμισμα».

Η δομή του blockchain που χρησιμοποιεί εκτός από τους χρήστες απαιτεί και την ύπαρξη εξορυκτών (miners). Οι τελευταίοι με τη χρήση εξειδικευμένου υλισμικού (hardware) δαπανούν υπολογιστική ισχύ για να επεξεργαστούν οι συναλλαγές, να ασφαλιστεί το δίκτυο και για να παραμείνουν όλοι στο σύστημα συγχρονισμένοι μαζί. Το σύστημα τους επιβραβεύει με αντάλλαγμα εκείνο τον miner που θα καταφέρει να λύσει πρώτος το block μέσω πολυσύνθετων μαθηματικών διεργασιών. Η ανταμοιβή είναι μεγάλη, $12.5 \text{ btc} \times 9000\$ = 112500\$$, για το λόγο αυτό οι miners συγκεντρώνονται σε ομάδες (pools) ώστε να έχουν μεγαλύτερη πιθανότητα να κερδίσουν το έπαθλο και μετά να το μοιραστούν ανάλογα με την επεξεργαστική ισχύ που προσέφερε ο καθένας.

Οι πρώτες προδιαγραφές (white paper) του Bitcoin και η απόδειξη της έννοιας (Proof of Work) του δημοσιεύθηκαν το 2009 από τον Satoshi Nakamoto. (<https://nakamotoinstitute.org/static/docs/bitcoin.pdf>). Ο δημιουργός του (ενδέχεται να είναι παραπάνω από ένα πρόσωπα) Satoshi αποσύρθηκε από το έργο αυτό στα τέλη του 2010 χωρίς να αποκαλύψει πολλά για τον εαυτό του. Από τότε, η κοινότητα του Bitcoin μεγάλωσε εκθετικά με πολλούς προγραμματιστές που ασχολούνται με το Bitcoin.

Η «ανωνυμία» του Satoshi δημιουργούσε αμφιβολίες για τα κίνητρα και το που αποσκοπούσε αυτό το εγχείρημα, μία άποψη που κέρδιζε έδαφος με παρανοήσεις γύρω από την ανοιχτού κώδικα φύση του Bitcoin. Το λογισμικό και το πρωτόκολλο του Bitcoin εκδίδονται ανοιχτά προς όλους και οποιοσδήποτε προγραμματιστής μπορεί να επιθεωρήσει τον κώδικα ή να φτιάξει τις δικές του τροποποιημένες εκδόσεις του λογισμικού Bitcoin. Ακριβώς όπως και με τους τωρινούς προγραμματιστές, η επιρροή του Satoshi περιορίστηκε στις αλλαγές που έκανε και υιοθετήθηκαν από άλλους και συνεπώς δεν είχε τον έλεγχο του Bitcoin. Έτσι, η ταυτότητα του εφευρέτη του Bitcoin δεν έχει αποκαλυφθεί μέχρι σήμερα.

Το Bitcoin αποτελεί ουσιαστικά ένα συναινετικό δίκτυο που παρέχει τη δυνατότητα ενός νέου συστήματος πληρωμών και μιας εντελώς ψηφιακής μορφής

χρημάτων. Πρόκειται για το πρώτο αποκεντρωμένο δίκτυο πληρωμής μεταξύ ομότιμων (peer-to-peer) που λειτουργεί από τους χρήστες του χωρίς κεντρική αρχή ή μεσάζοντες.

Το μπίτκοϊν (Bitcoin) αποτελεί στη βάση του ένα λογισμικό ανοιχτού κώδικα (open-source protocol). Κατά συνέπεια, ο πηγαίος κώδικας του λογισμικού είναι δημόσιος και διαθέσιμος σε όποιον επιθυμεί να ελέγξει τις λεπτομέρειες της λειτουργίας του. Η ανωτέρω αρχή επιτρέπει σε οποιονδήποτε την ελεύθερη και δωρεάν αντιγραφή και ανάπτυξη δικού του λογισμικού βασισμένου στο υπάρχον.

Το λογισμικό αποτελεί μία μέθοδο για την επίτευξη των παρακάτω κύριων στόχων:

- Θέσπιση κριτηρίων παραγωγής και συναλλαγής των ανταλλάξιμων μονάδων του λογισμικού (bitcoins),
- Διατήρηση των πληροφοριών ιδιοκτησίας των μονάδων των bitcoin που έχουν ήδη παραχθεί,
- Δυναμική επιβεβαίωση της εγκυρότητας των παραπάνω, χωρίς την ανάγκη ύπαρξης κεντρικής οντότητας ελέγχου, πιστοποίησης ή διακρίβωσης.

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.

Η ισχύς του δικτύου εξασφαλίζεται από την αποδοχή του από τους χρήστες. Το δίκτυο το οποίο αποτελούν οι χρήστες του Bitcoin, αποτελείται από χρήστες της ίδιας εκδοχής του λογισμικού. Αλλαγές στον κώδικα προτείνονται στην κοινότητα, αλλά η συναίνεση της κοινότητας των χρηστών και η αποδοχή τους είναι που δημιουργεί το δίκτυο.

Η πεποίθηση των χρηστών:

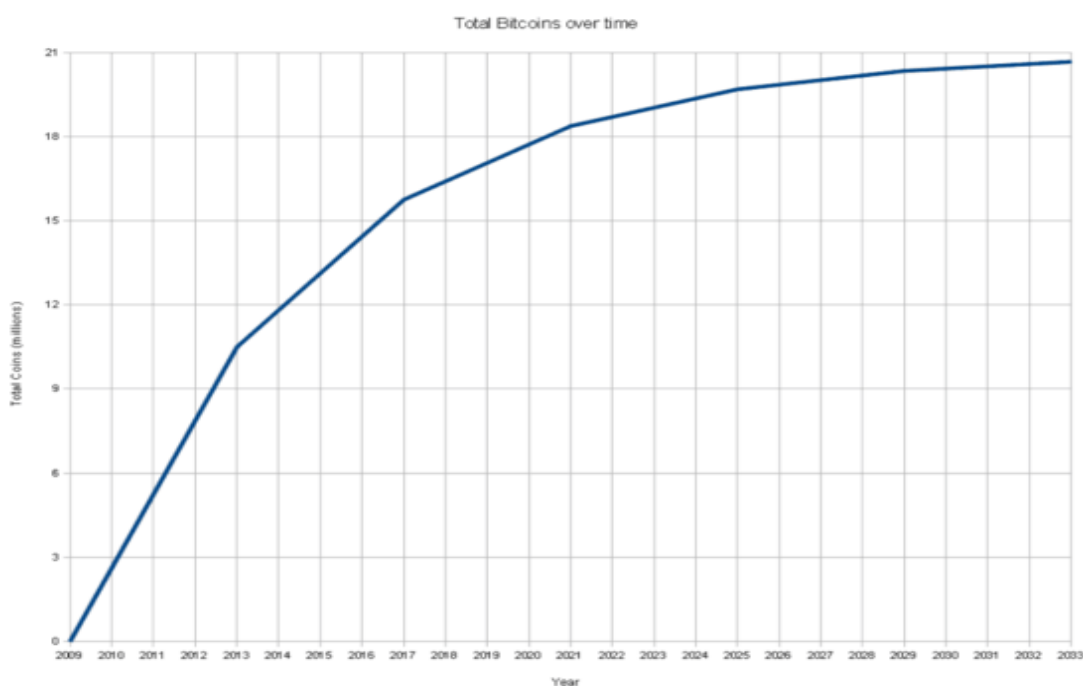
- στην διαφάνεια του πηγαίου κώδικα του λογισμικού,
- στην ακεραιότητα και διαφάνεια των συναλλασσόμενων πληροφοριών,
- στην στιβαρότητα του δικτύου από κακόβουλες επιθέσεις,

- στην προγραμματισμένα περιορισμένη παραγωγή bitcoins,
- στην προστασία που παρέχουν οι κρυπτογραφικοί αλγόριθμοι που χρησιμοποιούνται ενάντια σε κακόβουλη εκμετάλλευση του δικτύου.

Όπως και άλλοι συμπληρωματικοί λόγοι, είναι συνολικά υπεύθυνοι για την αποδοχή του από τους χρήστες, αλλά και την εξάπλωσή του σε νέους. Αυτό το λογισμικό και οι εξελίξεις του αποτελούν τον πυρήνα του συστήματος συναλλαγής bitcoins. Όποια αξία βρίσκουν οι χρήστες αποτυπώνεται αποκλειστικά στην αξία με την οποία είναι διατεθειμένοι να τα ανταλλάξουν, η οποία με τη σειρά της βασίζεται αποκλειστικά στους νόμους της προσφοράς και της ζήτησης, χωρίς ενδιάμεσα μέρη (χώρες, κεντρικούς εκδότες ή αρχές).

Για να μπορούν να είναι χρήσιμα σαν μέσο συναλλαγής πρέπει να εισάγονται στην κυκλοφορία σταδιακά για την κάλυψη των συναλλακτικών αναγκών, αλλά και να είναι πεπερασμένα σε συνολικό αριθμό. Αυτό επιτυγχάνεται τεχνητά και ο ρυθμός παραγωγής τους όπως και το μέγιστο πλήθος, αποτελούν μέρος των κανόνων του δικτύου.

Ο μέγιστος αριθμός που θα παραχθεί ποτέ είναι 21.000.000 και ο ρυθμός παραγωγής τους θα ελαττώνεται σταδιακά έως περίπου το 2140 οπότε και 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 παραγωγής bitcoin προσαρμόζεται τεχνητά ώστε να ακολουθεί περίπου την παρακάτω καμπύλη:



Εικόνα 3. Ρυθμός παραγωγής Bitcoin

Όπως κάθε τεχνολογική καινοτομία, έτσι και το Bitcoin έχει δύο όψεις. Τα βασικά πλεονεκτήματα συνοψίζονται στα εξής :

- **Ταχύτητα Συναλλαγών / Διεθνής Φύση:** Οι συναλλαγές σε bitcoin συμβαίνουν άμεσα και ενημερώνουν ταυτόχρονα όλο το δίκτυο παγκοσμίως. Αυτό δεν απαιτεί άλλες υποδομές πέρα από κάποια μορφή του δωρεάν λογισμικού σε υπολογιστή ή σε Smartphone, και σύνδεση στο διαδίκτυο.
- **Εξαιρετικά Χαμηλό κόστος συναλλαγών (low fees):** Το παρόν κόστος για κάθε συναλλαγή ανεξαρτήτως μεγέθους ανέρχεται περίπου στα 0.05\$ και είναι προαιρετικό, αν δεν υπάρχει βιασύνη επιβεβαίωσης της συναλλαγής. Το ποσό αυτό αποδίδεται αυτόματα στους χρήστες, που εκτελούν τους ελέγχους των συναλλαγών και την επιβεβαίωση της αντικειμενικότητάς του, ως αμοιβή για την επεξεργαστική ισχύ που επενδύουν στην προστασία του δικτύου από κακόβουλες επιθέσεις.
- **Έλεγχος από το χρήστη / Προστασία από υφαρπαγή:** Καθώς ο χρήστης είναι ο μόνος που έχει τη δυνατότητα να εκτελέσει συναλλαγές και εφόσον δεν έχει παραχωρήσει αυτό το δικαίωμα, και έχει προστατεύσει λογικά την πρόσβαση στα bitcoin του, είναι πρακτικά αδύνατο να κλαπούν ή να υφαρπαχτούν από τρίτους (εφόσον η κρυπτογράφηση δεν παραβιαστεί).
- **Φορητότητα:** Ανεξάρτητα από το πλήθος τους, τα bitcoins και τα «πορτοφόλια» αποθήκευσης ή οι κωδικού πρόσβασης σε αυτά είναι ουσιαστικά πάρα πολύ μικρά σε μέγεθος, και μπορούν να μεταφερθούν εύκολα, να καταγραφούν σε χαρτί, ακόμα και να απομνημονευτούν. Επίσης, κάτι αδύνατο για συμβατικές αξίες, μπορούν να αντιγραφούν ώστε να υπάρχουν αντίγραφα ασφαλείας σε περίπτωση καταστροφής των αρχικών.
- **Διαφάνεια Συναλλαγών:** Όλες οι συναλλαγές που έχουν εκτελεστεί ποτέ στο δίκτυο είναι δημόσια διαθέσιμες και διαφανείς. Έτσι, οποιοσδήποτε μπορεί να εξετάσει οποιαδήποτε διεύθυνση και να δει τις προηγούμενες συναλλαγές που έχουν εκτελεστεί με αυτήν, το πλήθος των bitcoin που έχουν μετακινηθεί, όπως και το που έχουν σταλεί.
- **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. Αντίστοιχου βεληνεκούς επιτυχημένα εγχειρήματα ανοιχτού λογισμικού αποτελούν το Linux όπως και το Bit torrent.

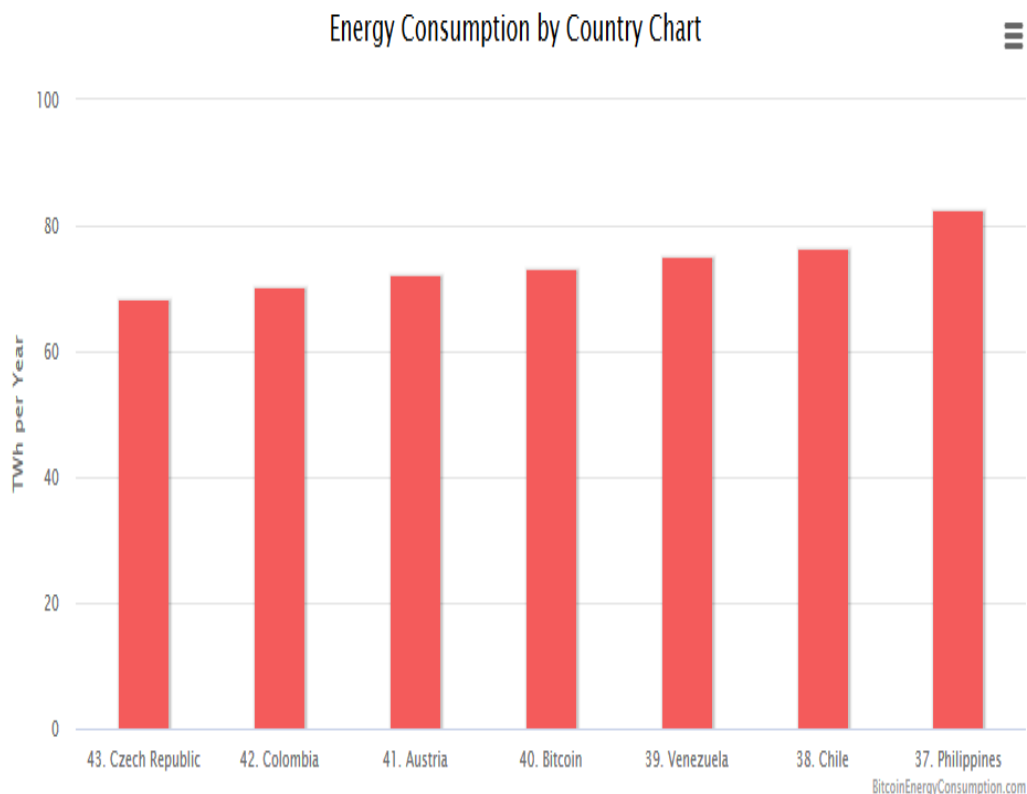
- **Αποκεντρωμένη Φύση (Decentralized):** Ένα από τα πιο σημαντικά χαρακτηριστικά του δικτύου, είναι η αποκεντρωμένη φύση του, που δεν απαιτεί καμία κεντρική αρχή ελέγχου ή επιβεβαίωσης. Κάθε κόμβος του δικτύου το ενισχύει περαιτέρω, αλλά αν προσβληθεί με κάποιο τρόπο, η λειτουργία του συνολικού δικτύου δεν επηρεάζεται ανάλογα. Ο μόνος τρόπος να σταματήσει να δουλεύει το δίκτυο είναι να αποκοπούν όλοι οι υπολογιστές του δικτύου μεταξύ τους. Ακόμα και τότε, με την επαναλειτουργία του διαδικτύου, το δίκτυο συνεχίζει ακριβώς εκεί που σταμάτησε.
- **Υποδιαιρέσεις:** Κάθε bitcoin είναι υποδιαιρέσιμο έως 8 δεκαδικά ψηφία (έως 0,00000001) που ονομάζονται Satoshi, επιτρέποντας μικρο-συναλλαγές(nano-transactions) που δεν είναι δυνατές με άλλα μέσα ή συμβατικά νομίσματα. Η προσθήκη περισσότερων ακόμα δεκαδικών επαφίεται στην συναίνεση του δικτύου αν αυτό χρειαστεί στο μέλλον.
- **Μη αντιστρέψιμη φύση:** Όλες οι συναλλαγές με bitcoin είναι τελικές και μη αντιστρέψιμες. 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. Αυτό συνήθως δίνει επιπλέον κίνητρα σε επιχειρήσεις να προσφέρουν τα προϊόντα τους σε χαμηλότερες τιμές, εξαιτίας της άμεσης και αμετάκλητης πληρωμής.
- **Ιδιωτικότητα συναλλαγών:** Κάθε χρήστης μπορεί να δημιουργήσει, μέσω του λογισμικού, σχεδόν απεριόριστο αριθμό διευθύνσεων μέσω των οποίων να εκτελέσει τις συναλλαγές του. Αυτές οι διευθύνσεις είναι ψευδώνυμες, δεν έχουν δηλαδή κάποια άμεση σχέση με τα πραγματικά στοιχεία ή την τοποθεσία του χρήστη, παρόλο που έχουν αναγνωρίσιμα χαρακτηριστικά ώστε να εντοπίζονται από το δίκτυο. Με αυτό τον τρόπο μπορεί ο χρήστης να διατηρήσει την ιδιωτικότητά του απομπλέκοντας τις συναλλαγές του από τα προσωπικά του στοιχεία. Αυτό δεν συνεπάγεται εξ' ορισμού ανωνυμία συναλλαγών καθώς όλες οι συναλλαγές δημοσιεύονται, και έστω και μία συναλλαγή να έχει γνωστό (δημόσιο) αποδέκτη, ίσως μπορεί να εξαχθεί από συμπληρωματικά στοιχεία η ταυτότητα του χρήστη.

Από την άλλη, η δημιουργία και η χρήση του Bitcoin ως μέσω συναλλαγής εγκυμονεί ορισμένους κινδύνους. Τα κύρια μειονεκτήματα είναι τα εξής:

- Απώλεια ιδιωτικών κλειδιών:** Το μόνο που ένας κακόβουλος χρήστης χρειάζεται ώστε να αποκτήσει έλεγχο των bitcoin του χρήστη, είναι η γνώση των ιδιωτικών κλειδιών του. Παρόλο που το λογισμικό ήδη παρέχει ικανή προστασία για το μέσο χρήστη, χρειάζεται εγρήγορση στην προστασία απέναντι σε ιούς ή κακόβουλο λογισμικό ή άλλου είδους παραβιάσεις (φυσικές ή ψηφιακές). Επιπρόσθετα, απώλεια πρόσβασης στα ιδιωτικά κλειδιά για άλλους λόγους (π.χ. θάνατος) καθιστά αδύνατη την ανάκτηση των bitcoins, χαρακτηριστικό παράδειγμα αποτελεί ο θάνατος του Διευθύνοντα Συμβούλου του QuadrigaCX exchange, με έδρα τον Καναδά ο οποίος απεβίωσε, όντας ο μοναδικός κάτοχος των ιδιωτικών κλειδιών με πρόσβαση σε 180 εκατομμύρια ευρώ σε bitcoin.
- Διακύμανση ισοτιμίας:** Καθώς τα bitcoin δεν έχουν κάποια κεντρική αρχή να παρεμβαίνει στις διακυμάνσεις στην προσφορά και την ζήτηση όπως συμβαίνει με για παράδειγμα με τα κρατικά νομίσματα, είναι ευαίσθητο σε μεγαλύτερες διακυμάνσεις της ισοτιμίας του με τα περισσότερα νομίσματα. Επιπλέον παράγοντας που επηρεάζει το παραπάνω φαινόμενο είναι η σχετικά μικρή κεφαλαιοποίηση (\$131.751.818.500, <https://coinmarketcap.com>, 11/12/2019) της αγοράς, πράγμα που σημαίνει ότι όταν συναλλάσσονται μεγάλοι όγκοι bitcoin, επηρεάζουν δυσανάλογα τις ισοτιμίες στα ανταλλακτήρια. Το φαινόμενο αυτό δείχνει να εξαλείφεται στην πάροδο του χρόνου καθώς η κεφαλαιοποίηση αυξάνεται με το χρόνο (λόγο αύξησης των διαθέσιμων bitcoin (circulating supply) αλλά κυρίως λόγω της ανάπτυξης κατάλληλων υποδομών και χρηματοοικονομικών προϊόντων που ήδη υφίστανται στις κλασσικές κεφαλαιαγορές, όπως Contract of Difference (CFDs) και Futures. Επιπλέον, η μεγάλη ταχύτητα συναλλαγών που προσφέρει το bitcoin, δίνει τη δυνατότητα για άμεσες δράσεις και αντιδράσεις στις αγορές, με αποτέλεσμα ορισμένες φορές ο παρορμητισμός των χρηστών/επενδυτών να προσαρμοστούν στις αλλαγές πυροδοτεί απότομες πτώσεις ή ανόδους στην τιμή.
- Ασαφές νομικό πλαίσιο:** Παρόλο που η ευρωπαϊκή νομοθεσία έχει λάβει μέτρα για τη θέσπιση όρων σε ότι αφορά κεντρικά ελεγχόμενα ή εκδιδόμενα ψηφιακά νομίσματα, η αποκεντρωμένη φύση των bitcoin, όπως και άλλα από τα χαρακτηριστικά τους, εισάγουν νέες παραμέτρους που δεν έχουν εξεταστεί σε όλο τους το εύρος ακόμα. Παρόλα αυτά τόσο στις Ηνωμένες Πολιτείες της Αμερικής, όσο και στην Ευρωπαϊκή Ένωση η κατοχή, συναλλαγή και η χρήση bitcoin θεωρείται νόμιμη και πλαισιώνεται από (https://www.supremecourt.gov/opinions/17pdf/17-530_6537.pdf) στις ΗΠΑ και από τον ευρύ καθορισμό ως "ιδιωτικά χρήματα" (Private money), από την Ευρωπαϊκή Κεντρική Τράπεζα. Ο

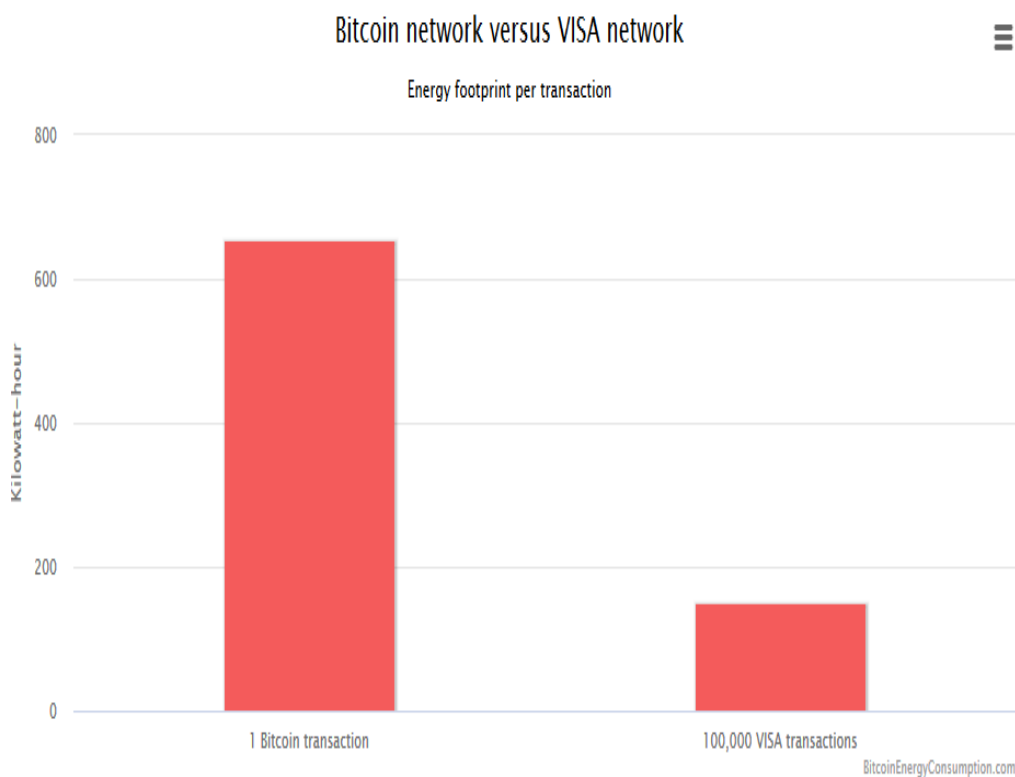
βασικός κίνδυνος προκύπτει από την έλλειψη ολοκληρωμένου, παγκόσμιου νομοθετικού πλαισίου που να προστατεύει τους χρήστες/ επενδυτές από κακόβουλους διαχειριστές συναλλακτηρίων αλλά και κάνει εφικτή την απόδοση δικαιοσύνης εκεί που χρειάζεται. Κατά καιρούς έχουν πραγματοποιηθεί αρκετές «ληστείες» σε μορφή κυβερνοεπίθεσης σε μεγάλα συναλλακτήρια , υφαρπάζοντας τεράστια ποσά , χωρίς να συλληφθούν ποτέ οι δράστες(MT GoX,2014,350million \$) αλλά και σχήματα Ponzi(BitConnect,2018, 2.6 billion \$). Το κύριο σημείο στο οποίο ασκείται κρατική επίβλεψη είναι το σημείο ανταλλαγής με τα κρατικά νομίσματα και ειδικότερα ότι έχει σχέση με τα «KnowYourClient» και «AntiMoneyLaundering» νομικά πλαίσια τοπικά και διεθνώς.

- **Κατανάλωση Ηλεκτρικής Ενέργειας:** Ένα όχι και τόσο εμφανές πρόβλημα αποτελεί η εκτεταμένη ανάγκη για ηλεκτρικό ρεύμα. Αυτό προκύπτει από το γεγονός ότι η εξόρυξη κρυπτονομισμάτων απαιτεί υψηλή διάθεση επεξεργαστικής ισχύς και αυτή με τη σειρά της μεγάλα ποσά ενέργειας. Αξίζει να σημειωθεί ότι κατά κύριο λόγο η εξόρυξη γίνεται με χρήση καρτών γραφικών(GPU) ή εξειδικευμένων μικροϋπολογιστών (dedicated hardware) , τα οποία έχουν απαιτήσεις σε άεργο και όχι ενεργό ισχύ. Το μέγεθος του προβλήματος που προκύπτει γίνεται αντιληπτό από το παρακάτω διάγραμμα σε σύγκριση της κατανάλωσης ενέργειας με άλλα κράτη.



Εικόνα 4. Σύγκριση κατανάλωσης ενέργειας μεταξύ κρατών (Πηγή: BitcoinEnergyConsumption.com)

Αλλά και με ανταγωνιστές παλαιότερων τεχνολογιών :



Εικόνα 5. Σύγκριση Bitcoin network με VISA network (Πηγή: BitcoinEnergyConsumption.com)

Επιπλέον, αξίζει να αναφερθεί ότι όσο περισσότεροι εξορυκτές συμμετέχουν στη διαδικασία ή όσο αυξάνουν την επεξεργαστική τους ισχύ για να κερδίσουν μεγαλύτερο μερίδιο στην ανταμοιβή, τόσο αυξάνεται και η απαιτούμενη ενέργεια. Πιθανολογείται από μακροοικονομική άποψη ότι το κόστος ενέργειας θα είναι ο κύριος παράγοντας που μπορεί να θέσει ανώτατο όριο στην άνοδο της τιμής του Bitcoin.

1.3.2. Ethereum

Το «εθέριουμ» (Ethereum) είναι μια ανοιχτή πλατφόρμα κατανεμημένων υπολογιστών, δηλαδή μια δημόσια πλατφόρμα που υποστηρίζει την λειτουργία των έξυπνων συμβολαίων (smart contracts). Παρέχει μια αποκεντρωμένη εικονική μηχανή Turing-complete, την Εικονική Μηχανή Ethereum (EVM), η οποία μπορεί να εκτελέσει τα Smart Contracts μέσω ενός δικτύου δημόσιων κόμβων. Το Ethereum παρέχει επίσης ένα κρυπτονόμισμα που ονομάζεται "Ether", το οποίο μπορεί να μεταφερθεί μεταξύ λογαριασμών και να χρησιμοποιηθεί για την αντιστάθμιση των συμμετεχόντων

κόμβων για υπολογισμούς που εκτελούνται. Το «Ether» αποτελεί το δεύτερο, σε μέγεθος βάση κεφαλαιοποίησης και ημερήσιου όγκου συναλλαγών, κρυπτονόμισμα. Επιπρόσθετα, το "Gas", ένας μηχανισμός εσωτερικής τιμολόγησης συναλλαγών, χρησιμοποιείται για την αντιμετώπιση των ανεπιθύμητων μηνυμάτων και την κατανομή πόρων στο δίκτυο. Ουσιαστικά, το Ethereum είναι μια παγκόσμια υπολογιστική μηχανή.

Το Ethereum προτάθηκε στα τέλη του 2013 από τον Vitalik Buterin (Ethereum: A next-generation smart contract and decentralized application platform, 2013, V. Buterin), έναν ερευνητή και προγραμματιστή κρυπτογράφησης. Η ανάπτυξη και η υλοποίηση χρηματοδοτήθηκε από μία διαδικτυακή κοινότητα και το blockchain τέθηκε σε λειτουργία στις 30 Ιουλίου 2015, με 11,9 εκατομμύρια νομίσματα "premined" για την κοινότητα που το υποστήριξε. Μια από τις ιδέες που βρήκαν εφαρμογή στην πλατφόρμα του Ethereum, ήταν η δημιουργία ενός ανεξάρτητου και αυτόνομου οργανισμού με ψηφιακά κεφάλαια επιχειρηματικών συμμετοχών (venture capital funds). Ο σκοπός αυτών των οργανισμών είναι να υποστηρίζουν οικονομικά νέες startup επιχειρήσεις. Τον Απρίλιο του 2016 ο προγραμματιστής Christoph Jentzsch ανακοίνωσε τον οργανισμό DAO. Τα αρχικά σημαίνουν Decentralized Autonomous Organization, ή αλλιώς Αποκεντρωμένη Αυτόνομη Οργάνωση. Με την τεχνολογία των smart contracts του Ethereum, ο Jentzsch δημιούργησε τα νομίσματα DAO tokens. Μέσω της διαδικασίας του Initial Coin Offering (ICO) τα προσφέρει για αγορά σε υποψήφιους επενδυτές. Το όνομα ICO είναι το αντίστοιχο του IPO (Initial Public Offering), που είναι η διαδικασία όταν μια εταιρεία μπαίνει στο χρηματιστήριο και διαθέτει μετοχές για αγορά από το κοινό.

Με το ICO, κάθε μέλος αγοράζει ένα ποσοστό των DAO tokens και συμμετέχει στην αύξηση του κεφαλαίου του οργανισμού. Επιπλέον, έχει το δικαίωμα της ψήφου για την τελική επιλογή των startups που έχουν αιτηθεί χρηματοδότησης. Εάν οι νέες επιχειρήσεις πετύχουν, ένα μέρος από τα κέρδη τους επιστρέφεται με την μορφή μερίσματος στους επενδυτές, ανάλογα με το ποσοστό των DAO tokens που κατέχουν. Τον Μάιο του 2016, η αξία των χρημάτων που επενδύθηκαν σε DAO tokens ξεπερνούσε τα \$150 εκατομμύρια, προσελκύοντας πάνω από 11.000 επενδυτές.

Στις αρχές του 2018, με την ιλιγγιώδη αύξηση στην αξία του Ethereum, σημαντικός αριθμός χρηστών διεθνώς ενδιαφέρεται για το mining. Ενώ το mining του Ethereum είναι αντικειμενικά ευκολότερο από του Bitcoin, και σε αυτή την περίπτωση έχουμε ξεπεράσει κατά πολύ τον απλό υπολογιστή. Ένα έμπρακτο παράδειγμα είναι ο Anthony Garreffa, ένας από τους συντελεστές της ιστοσελίδας Tweaktown. Στα μέσα Ιουνίου 2017 είχε κέρδος \$900 το μήνα σε. Όμως ο εξοπλισμός που χρησιμοποιεί για το mining είναι αξίας \$10.000.

Το Ethereum χρησιμοποιείται πέρα της ανταλλακτικής του αξίας (<https://www.coinbase.com/charts>) και σαν «φόρος» για αλληλεπίδραση με τα Smart

Contract στο δίκτυο αλλά και ως αξία κοστολόγησης για την υπολογιστική ισχύ που χρειάζεται το Ethereum δημόσιο δίκτυο blockchain.

1.3.2.1. Έξυπνα συμβάσεις (Smart Contracts)

Το 1994, ο Nick Szabo, νομικός μελετητής και κρυπτογράφος, συνειδητοποίησε ότι οι αποκεντρωμένες δομές θα μπορούσαν να χρησιμοποιηθούν για έξυπνες συμβάσεις, διαφορετικά αποκαλούμενες συμβάσεις αυτοελέγχου, αποκλειστικές συμβάσεις ή ψηφιακές συμβάσεις (smart contracts). Σε αυτή τη μορφή, οι συμβάσεις θα μπορούσαν να μετατραπούν σε κώδικα υπολογιστή, να αποθηκευτούν και να αναπαραχθούν στο σύστημα και να εποπτευθούν από το δίκτυο των υπολογιστών που τρέχουν το blockchain. Αυτό θα οδηγήσει επίσης σε ανατροφοδότηση μέσω blockchain, όπως η μεταφορά χρημάτων και η παραλαβή του προϊόντος ή της υπηρεσίας. Οι έξυπνες συμβάσεις (Smart Contracts) εξυπηρετούν στην συναλλαγή χρημάτων, ακινήτων, μετοχών και οποιουδήποτε αγαθού ή υπηρεσίας με διαφανή τρόπο, αποφεύγοντας παράλληλα τις υπηρεσίες κάποιου μεσάζοντα. Πιο συγκεκριμένα, οι έξυπνες συμβάσεις όχι μόνο καθορίζουν τους κανόνες και τις κυρώσεις γύρω από μια συμφωνία με τον ίδιο τρόπο που εφαρμόζει μια παραδοσιακή σύμβαση, αλλά και αυτομάτως επιβάλλουν αυτές τις υποχρεώσεις. Μια νομική δέσμευση-ψηφιακή συμφωνία με την δυνατότητα να αυτοπροσαρμόζεται δυναμικά, ανάλογα με τους όρους και τις συνθήκες της αγοράς ή του περιεχομένου που ενστερνίζεται, μέσω της υλοποίησης του συνόλου των ενσωματωμένων κανόνων από όπου χαρακτηρίζεται. Οι στόχοι των έξυπνων συμβολαίων είναι:

- η αφαίρεση της χειρωνακτικής παρέμβασης και της επιτήρησης
- η μείωση των σχετικών τελών, νομικών εξόδων και διαδικασιών
- ταχύτητα της δημιουργίας συμβολαίων και της εκτέλεσης του
- η αυτόματη μεταφορά κεφαλαίων μέσω αναγνωρίσιμων γεγονότων
- ευελιξία και εξελισσιμότητα των συμβολαίων

Συνοψίζοντας, τα κρυπτονομίσματα και κατ' επέκταση το blockchain αποτελεί σημείο σταθμό τόσο για την επιστήμη των υπολογιστών όσο και για τον κλάδο της οικονομίας. Ο συγκερασμός των δύο αυτών κλάδων έδωσε μία καινοτομία που θα διαδραματίσει κεντρικό ρόλο τόσο στον τομέα των χρηματοοικονομικών αλλά και στην καθημερινή ζωή των ανθρώπων. Στη συνέχεια θα παρουσιαστεί άλλη μία τεχνολογική εφαρμογή της πληροφορικής στα χρηματοοικονομικά, στο πλαίσιο της οποίας ανήκει και η μελέτη της παρούσας διπλωματικής εργασίας.

1.4. Αλγοριθμικές Συναλλαγές

Οι αλγοριθμικές συναλλαγές (algorithmic trading) ή αλλιώς αυτοματοποιημένες συναλλαγές (automated trading), black-box trading ή algo trading είναι εκείνες οι συναλλαγές που εκτελούνται από προκαθορισμένη σειρά εντολών(αλγόριθμος) με τη χρήση ηλεκτρονικών υπολογιστών. Η ύπαρξη τους εκπληγάζει από το γεγονός ότι εξέλιξη των υπολογιστών, του διαδικτύου και οι δυνατότητες που αυτά παρέχουν, κατέστησαν εφικτή την υλοποίηση κερδοφόρων συναλλαγών σε ταχύτητα και συχνότητα που είναι πρακτικά αδύνατο να διεκπεραιωθούν χειροκίνητα. Αναλυτικότερα, εξειδικευμένοι προγραμματιστές, μαθηματικοί, μηχανικοί και χρηματιστές συνδυάζουν τις γνώσεις τους για τη δημιουργία αλγορίθμων βασισμένους στην τιμή, την ποσότητα, το timing ή οποιαδήποτε μαθηματικό μοντέλο.

Οι μεσοπρόθεσμοι ή μακροπρόθεσμοι επενδυτές ή οι εταιρείες από την πλευρά της αγοράς - συνταξιοδοτικά ταμεία, αμοιβαία κεφάλαια(mutual funds), ασφαλιστικές εταιρείες - χρησιμοποιούν το algo-trading για να αγοράζουν αποθέματα σε μεγάλες ποσότητες όταν δεν θέλουν να επηρεάσουν τις τιμές των μετοχών με διακριτές επενδύσεις μεγάλου όγκου. Από την άλλη μεριά, οι βραχυπρόθεσμοι επενδυτές και οι market makers - διαμορφωτές της αγοράς (όπως επενδυτικές εταιρείες και brokerage houses), κερδοσκόποι και arbitrageurs- επωφελούνται από την αυτόματη εκτέλεση του συναλλαγών, μέσω του High Frequency Trading(HFT) και την παροχή ρευστότητας στην αγορά. Επίσης, συστηματικοί επενδυτές όπως hedge funds ή traders που κλείνουν long positions με μία short σε επενδυτικά προϊόντα τύπου exchange-traded funds (ETFs) ή forex, αξιοποιούν τις αλγοριθμικές συναλλαγές για να κάνουν πολύ αποτελεσματικότερο προγραμματισμό των θέσεων τους και αυτές να εκτελούνται αυτοματοποιημένα.

1.4.1. Στρατηγικές Αλγοριθμικών Συναλλαγών

Όλες οι στρατηγικές αλγοριθμικών συναλλαγών (Algorithmic Trading Strategies) βασίζονται στην εύρεση μίας ευκαιρίας-συναλλαγής η οποία επιτυγχάνει αύξηση των κερδών ή κάποια μείωση στα κόστη της εταιρείας. Οι πιο διαδεδομένες τέτοιες στρατηγικές είναι οι εξής :

1.4.1.1. Ευκαιρίες Arbitrage

Πρόκειται για επενδυτικά αγαθά που είναι εισηγμένα σε περισσότερα από ένα συναλλακτήρια, για παράδειγμα μία μετοχή που είναι εισηγμένη σε 2 χρηματιστήρια (dual-listed stock), και την ίδια χρονική στιγμή στο ένα πωλείται σε υψηλότερη τιμή από ότι αγοράζεται στο άλλο, δίνοντας έτσι τη δυνατότητα εκτέλεσης κερδοφόρας συναλλαγής με ουσιαστικά μηδενικό ρίσκο. Αυτό έγκειται στο γεγονός ότι τα 2 διαφορετικά συναλλακτήρια, κυρίως λόγω διαφοράς ώρας, έχουν οριακά διαφορετική προσφορά και ζήτηση σε ένα επενδυτικό αγαθό με αποτέλεσμα να διαμορφώνονται στιγμιαία διαφορετικές τιμές. Αξίζει να σημειωθεί εδώ ότι στην αγορά των κρυπτονομισμάτων, δεδομένου ότι η αγορά είναι αποκεντροποιημένη και ότι σχεδόν

όλα τα εμπορεύσιμα ζευγάρια συναλλαγών(trading-pairs) υπάρχουν σε όλα τα συναλλακτήρια, οι ευκαιρίες arbitrage είναι πολύ σύνηθες φαινόμενο. Επεκτείνοντας αυτό ένα βήμα πιο κάτω, είναι εύλογο να διατυπωθεί ότι οι arbitrageurs είναι εκείνοι που κρατούν ουσιαστικά κοντά τις τιμές των κρυπτονομισμάτων ανά τα συναλλακτήρια παγκοσμίως.

1.4.1.2. Στρατηγικές Παρακολούθησης Τάσεων(Trend-following)

Οι στρατηγικές παρακολούθησης τάσεων είναι οι πιο συνηθισμένες αλγοριθμικές στρατηγικές διαπραγμάτευσης που ακολουθούν τις τάσεις στους κινούμενους μέσους όρους, τα διαλείμματα καναλιών(channel breakouts), τις κινήσεις επιπέδων τιμών και τους συναφείς τεχνικούς δείκτες. Πρόκειται για τις πιο εύχρηστες και απλές στρατηγικές για την εφαρμογή μέσω αλγοριθμικών συναλλαγών επειδή αυτές οι στρατηγικές δεν συνεπάγονται χρήση πολύπλοκων αλγορίθμων για κατανόηση της αγοράς ή πρόβλεψη τιμών. Οι συναλλαγές ξεκινούν με βάση την εμφάνιση επιθυμητών τάσεων, οι οποίες είναι εύκολο και απλό να εφαρμοστούν χωρίς να εισέλθουν στην πολυπλοκότητα της προβλεπτικής ανάλυσης. Για παράδειγμα, η χρήση κινητών μέσων 50 και 200 ημερών είναι μια δημοφιλής στρατηγική που ακολουθεί τις τάσεις και χρησιμοποιείται ευρέως για την πρώτη εικόνα του επενδυτικού αγαθού.

1.4.1.3. Στρατηγικές βασισμένες σε μαθηματικά μοντέλα

Αποδεδειγμένα μαθηματικά μοντέλα, όπως η στρατηγική διαπραγμάτευσης Delta-Neutral, επιτρέπουν τη διαπραγμάτευση ενός συνδυασμού επιλογών επενδυτικών προϊόντων και της υποκείμενης ασφάλειας τους. Για παράδειγμα στην αγορά κρυπτονομισμάτων μπορεί ο επενδυτής να έχει στη κατοχή του Ethereums(Eth) ή future-contracts αυτών και αντίστοιχα short-positions σε CFDs Ethereums. Η Delta neutral, πιο συγκεκριμένα, είναι μια στρατηγική βελτιστοποίησης χαρτοφυλακίου που αποτελείται από πολλαπλές θέσεις με αντισταθμιστικό θετικό και αρνητικό δείκτη deltas, με σκοπό το άθροισμα τους να είναι όσο δυνατόν πιο κοντά στο 0. Ως delta ορίζεται ένας από 4 βασικούς δείκτες για αξιολόγηση του ρίσκου και εκφράζει το βαθμό της ευαισθησίας που έχει ένα παράγωγο στην κίνηση της τιμής του αγαθού. Το εύρος τιμών του δείκτη περιορίζεται από 1.0 έως -1.0, όπου θετικός είναι για αγορά ενός call και αρνητικός για put.

1.4.1.4. Αντιστροφή μέσου όρου (Mean Reversion)

Η στρατηγική αντιστροφής του μέσου όρου βασίζεται στην έννοια ότι οι υψηλές και οι χαμηλές τιμές ενός επενδυτικού στοιχείου είναι ένα προσωρινό φαινόμενο που αντιστρέφεται περιοδικά γύρω από τη μέση αξία τους (μέση τιμή). Ο εντοπισμός και ο καθορισμός του εύρους τιμών και η σχεδίαση και εφαρμογή του κατάλληλου αλγορίθμου βασισμένου σε αυτό, επιτρέπει την αυτόματη τοποθέτηση συναλλαγών όταν η τιμή του ζητούμενου επενδυτικού αγαθού διαρρηγνύεται εντός και εκτός του καθορισμένου εύρους τιμών.

Στρατηγική σταθμισμένων μέσων όρων (Volume/ Time Weighted Average Price)

Η στρατηγική σταθμισμένων μέσων όρων έγκειται στη μεθοδολογία κατακερματισμού μίας συναλλαγματικής εντολής με μεγάλο όγκο σε αρκετές μικρότερες, ώστε δυναμικά να συναλλάσσονται μικρότερα κομμάτια του όγκου αυτού και να μην επηρεάσει η ίδια η συναλλαγή τα ποσοτικά χαρακτηριστικά που ενδιαφέρουν τον επενδυτή.

Είναι χρήσιμο να διευκρινιστεί ότι μία εντολή συναλλαγής επενδυτικού αγαθού με μεγάλο όγκο επηρεάζει την τιμή του, καθώς παρεμβαίνει ακαριαία είτε στην προσφορά, είτε στη ζήτηση του προϊόντος αυτού. Στο πλαίσιο αυτό χρησιμοποιείται η στρατηγική των σταθμισμένης τιμής βάση όγκου (VWAP) εστιάζοντας σε συγκεκριμένα ιστορικά προφίλ του όγκου συναλλαγών, ώστε να αποφανθεί ο κατάλληλος όγκος ανά συναλλαγή που δεν επηρεάζει αισθητά την τιμή. Επιπλέον, αξιοποιείται και η στρατηγική χρονικά σταθμισμένων μέσων τιμών (TWAP) (<https://okexsupport.zendesk.com/hc/en-us/articles/360027486392-Spot-Trading-What-is-Time-weighted-average-price-TWAP->) χρησιμοποιώντας ομοιόμορφα διαιρεμένα χρονικά διαστήματα μεταξύ χρόνου έναρξης και λήξης.

Στόχος είναι να εκτελεστεί η εντολή κοντά στη μέση τιμή μεταξύ των χρόνων έναρξης και λήξης, ελαχιστοποιώντας έτσι τις επιπτώσεις στην αγορά.

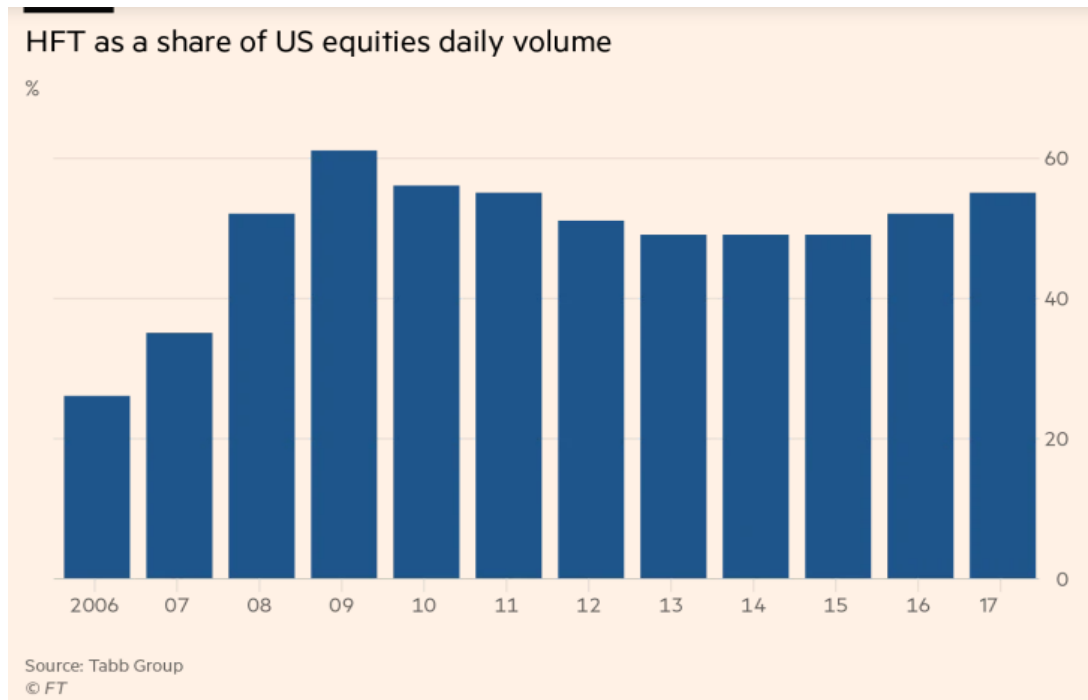
1.4.2. Συναλλαγές υψηλής συχνότητας

Οι συναλλαγές υψηλής συχνότητας (High Frequency Trading - HFT) είναι μία μέθοδος διεκπεραίωσης συναλλαγών από πολύ ισχυρά υπολογίσιμα προγράμματα σε κλάσματα του δευτερολέπτου. Χρησιμοποιούνται αλγόριθμοι για εκτενής ανάλυση της αγοράς και πραγματοποιούνται οι συναλλαγές βάση της κατάστασης που βρίσκεται η αγορά. Αποτελούν ουσιαστικά μία υπο-ενότητα του ευρύτερου algorithmic trading που περιλαμβάνει το μεγαλύτερο ποσοστό αυτού, καθώς όπως είναι εμφανές οι περισσότερες στρατηγικές αλγοριθμικών συναλλαγών απαιτούν την εκτέλεση τους σε πολύ υψηλές συχνότητες.

Αξίζει να σημειωθεί, ότι εκτός από κερδοφόρες συναλλαγές, το high frequency trading συνεισφέρει ουσιαστικά στην ρευστότητα τις αγοράς καθώς κάθε χρονική στιγμή εκτελείται πολύ μεγάλο πλήθος συναλλαγών, σε σχέση με τις παραδοσιακές εντολές που μπορεί να χρειαστούν αρκετές μέρες να εκτελεστούν. Η κυριαρχία των συναλλαγών υψηλής συχνότητας στην αγορά ξεκίνησε κυρίως μετά την κρίση στο χρηματοπιστωτικό κλάδο και την κατάρρευση της Lehman Brothers το 2008, καθώς λόγω έλλειψης ρευστότητας, τα χρηματιστήρια αρχίσαν προσφέρουν αντάλλαγμα σε εκείνους τους trader που εκτελούσαν εύλογο όγκο ημερήσιων συναλλαγών και παρείχαν ρευστότητα στην αγορά. Χαρακτηριστικό παράδειγμα αποτελεί το Πρόγραμμα Ρευστότητας (Supplemental Liquidity Providers (SLPs) του New York Stock Exchange (<https://www.nyse.com/markets/liquidity-programs>) που υποστηρίζεται και

εξελίσσεται με την συμβολή των μεγαλύτερων ονομάτων της Wall Street , όπως Goldman Sachs & Company, Virtu Americas LLC , Hudson River Trading , Citadel Securities LLC κλπ.

Στη συνέχεια, παρατίθεται ραβδόγραμμα των HFT σε ποσοστό των ημερήσιων συναλλαγών μετοχών στις Η.Π.Α.

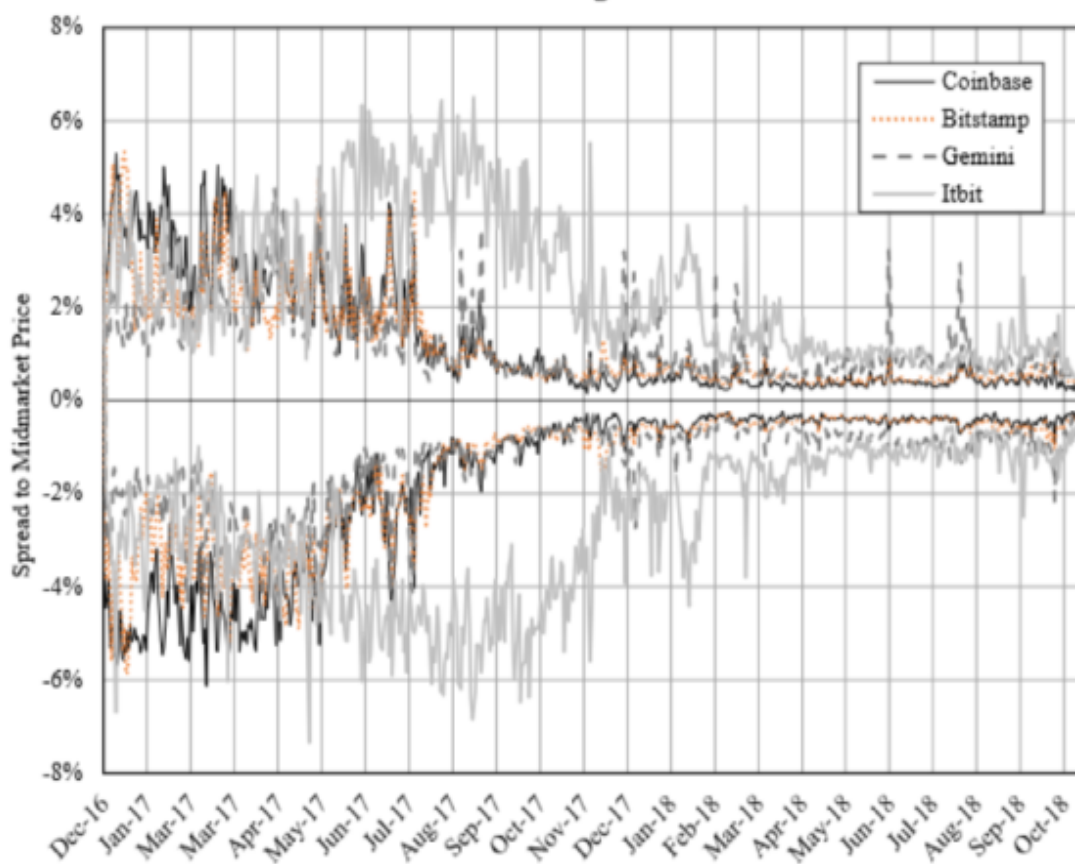


Εικόνα 6. HFT σε ποσοστό των ημερήσιων συναλλαγών μετοχών στις Η.Π.Α (Πηγή: Financial Times)

Στην αγορά κρυπτονομισμάτων κυρίως με την εισαγωγή της μόχλευσης (leverage) τόσο στα options και futures αλλά και στην ίδια τα κρυπτονομίσματα, η ανάγκη για ρευστότητα στα συναλλακτήρια είναι μεγάλη. Με γνώμονα αυτό, τα μεγαλύτερα συναλλακτήρια διαθέτουν εκτεταμένα τόσο προγράμματα επιβράβευσης για παροχή ρευστότητας μέσω συναλλαγών πολύ υψηλής συχνότητας όσο και τη δυνατότητα να δημιουργηθεί και επεξεργασθεί σε open source μορφή κώδικας για την αξιοποίηση και ανάπτυξη μεθόδων algorithmic trading.

Για να γίνει καλύτερα αντιληπτό το αντίκτυπο του High-Frequency-Trading στην αγορά των κρυπτονομισμάτων, παρατίθεται ένα διάγραμμα του κόστους να εκτελεσθεί μία εντολή 500.000\$ ή καλύτερα το spread του κρυπτονομίσματος για 4 από τα μεγαλύτερα συναλλακτήρια. Όπως είναι εμφανές με την είσοδο και την ανάπτυξη του στην αγορά το εύρος διαφοράς bid και ask έχει αμβλυνθεί αισθητά από περίπου 4% - 5% σε >0.8%.

Exhibit 17: Cost to Execute a \$500,000 Market Order at Eligible Exchanges



A bid and an ask line are shown for each exchange. Bid prices are negative numbers and ask prices are positive numbers.
Source: NYDIG analysis of exchange data

Εικόνα 7. Κόστος εκτέλεσης εντολής 500.000\$

2. Εισαγωγή στις προβλέψεις

2.1. Γενικά για τις προβλέψεις

Η επιστήμη των προβλέψεων έχει ως βασικό αντικείμενο ανάλυσης το ιστορικό(παρελθόν) ενός μεγέθους ή αγαθού και την επεξεργασία αυτού με τέτοιο τρόπο, ώστε να μπορεί να προσδιοριστεί ή να περιγραφεί η συμπεριφορά στοιχείων όπως το ρίσκο και η αβεβαιότητα για τις μελλοντικές εξελίξεις. Ο κλάδος των προβλέψεων θα μπορούσε να χαρακτηριστεί ως ένας συνδυασμός στατιστικής ανάλυσης δεδομένων, οικονομετρίας και επιχειρησιακής έρευνας. Πάρα τις απαιτούμενες γνώσεις για ποσοτική ανάλυση, οι προβλέψεις αποτελούν αναπόσπαστο κομμάτι της καθημερινότητας μας. Η παραγωγή μιας πρόβλεψης για την κάλυψη των εβδομαδιαίων αναγκών ενός νοικοκυριού ή ακόμα και η παραγωγή μιας πρόβλεψης για την τοποθέτηση ενός στοιχήματος ποδοσφαιρικού αγώνα είναι αποτελέσματα ανάλυσης εμπειρικών στοιχείων και δεδομένων που έχουμε χρησιμοποιήσει για να φτάσουμε σε μία εκτίμηση αποτελέσματος.

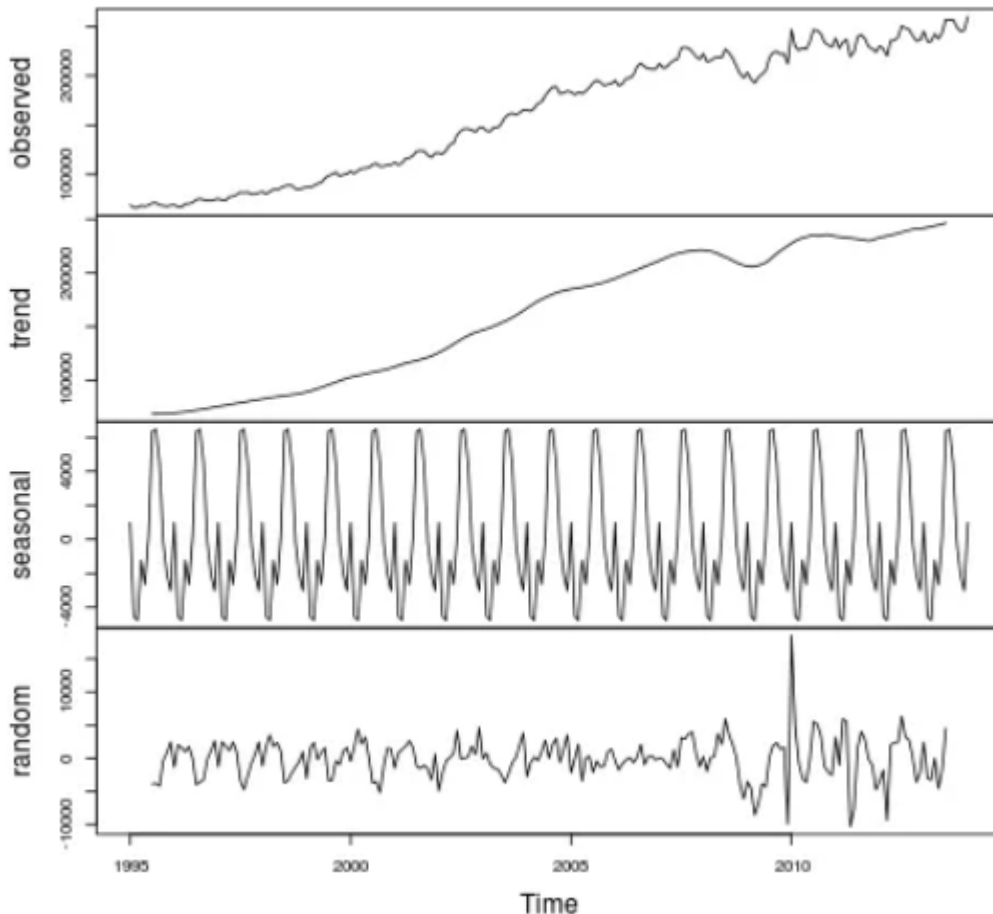
Οι προβλέψεις παρουσιάζονται σχεδόν σε όλες τις εκφάνσεις τόσο του επιστημονικού όσο και του επιχειρησιακού τομέα. Στα πλαίσια της επιχείρησης, οι αποφάσεις που καλείται να λαμβάνει ανά τακτά διαστήματα για τα μελλοντικά της σχέδια, είτε αυτά αφορούν κάποια επέκταση ή κάποια νέα επένδυση, υπόκεινται στην ανάγκη παραγωγής προβλέψεων. Η αποτελεσματικότητα των αποφάσεων αυτών είναι άρρηκτα συνδεδεμένη με την ικανότητα παραγωγής ακριβών και εύστοχων προβλέψεων. Για παράδειγμα, τα καταστήματα λιανικής καθορίζουν βάσει των δεδομένων που έχουν και της εμπειρίας του ιδιοκτήτη, το εκτιμώμενο ύψος πωλήσεων για κάποια χρονική περίοδο ώστε να πραγματοποιήσουν έγκαιρα και με ακρίβεια νέα προϊόντα για να καλύψουν τη ζήτησή τους επαρκώς, αντίστοιχα τα εργοστάσια έχουν την ανάγκη να προβλέψουν τη ζήτηση για να ρυθμίσουν την παραγωγή τους.

Αρκετοί χρηματοοικονομικοί αναλυτές και σύμβουλοι επενδύσεων βασίζονται στην παραγωγή ακριβών προβλέψεων για να προχωρήσουν σε επενδύσεις σε μετοχές και να αυτοματοποιήσουν την διαδικασία λήψης αποφάσεων σε συνθήκες με μεγάλη αβεβαιότητα. Σαφώς πέρα από τον επιχειρηματικό τομέα, η επιστήμη των τεχνικών προβλέψεων εφαρμόζεται άμεσα και σε άλλες επιστήμες, με διαφορετικές ίσως προσεγγίσεις και τεχνικές από αυτές που αναλύονται παρακάτω. Μεγάλο μέρος της δουλειάς των μετεωρολόγων έγκειται στη βραχυχρόνια ή μεσοπρόθεσμη πρόγνωση του καιρού, οι συντηρητές του Συστήματος Ηλεκτρικής Ενέργειας της χώρας έχουν ανάγκη την πρόβλεψη του πλήθους των ετήσιων βλαβών που θα προκύψουν ενώ το ίδιο το κράτος παράγει κάθε χρόνο προβλέψεις για μακροοικονομικούς δείκτες όπως το ΑΕΠ και ο πληθωρισμός, αλλά ακόμη και για την ενεργειακή κατανάλωση της χώρας. Στην παρούσα εργασία αυτή θα αναλυθεί κατά πόσο ο κλάδος των τεχνικών προβλέψεων μπορεί να εφαρμοστεί στην εύστοχη και ακριβής εκτίμηση της βραχυπρόθεσμης τιμής κλεισίματος στα συναλλακτήρια κρυπτονομισμάτων.

Λόγω του ευρύτατου φάσματος εφαρμογών, οι τεχνικές προβλέψεων έχουν κατηγοριοποιηθεί σε ποσοτικές και ποιοτικές. Ως ποσοτικές μεθόδους ορίζουμε αυτές που δίνουν αποτελέσματα αξιοποιώντας και πραγματοποιώντας υπολογισμούς σε τιμές του παρελθόντος (χρονοσειρές). Αντίθετα, ως ποιοτικές ονομάζουμε τα αποτελέσματα των εκτιμήσεων που παράγονται κυρίως από φυσικά πρόσωπα σε περιπτώσεις ανεπαρκών στοιχείων και βασίζονται κυρίως στην εμπειρία, τη διαίσθηση και τις ικανότητες των ανθρώπων που καλούνται να παράξουν εκτίμηση γινη πρόβλεψη. Οι ποιοτικές μέθοδοι χρησιμοποιούνται συνήθως σε συνδυασμό με κάποιες ποσοτικές.

2.2. Χαρακτηριστικά Χρονοσειρών

Βασικό στοιχείο για την επιστήμη των προβλέψεων αποτελεί η χρονοσειρά. Ως χρονοσειρά ορίζουμε ένα σύνολο από διαδοχικές παρατηρήσεις κάποιου φυσικού ή άλλου μεγέθους που δεν είναι ανεξάρτητες μεταξύ τους. Λόγω της εξάρτησης αυτής και μέσω των τεχνικών προβλέψεων, οι μελλοντικές τιμές της κάθε χρονοσειράς μπορούν να προσδιοριστούν ακριβώς με ντετερμινιστικό τρόπο. Φυσικά στις πραγματικές χρονοσειρές υπάρχει το στοιχείο της τυχαιότητας (ή τυχαίου παράγοντα) που αντιπροσωπεύεται από μια στατιστική μεταβλητή. Θεωρούμε ότι οι χρονοσειρές πραγματικών μεγεθών αντιπροσωπεύουν στοχαστικές διαδικασίες. Η ανάλυση των χρονοσειρών αυτών αποτελεί το βασικότερο αντικείμενο των προβλέψεων. Οι κλασσικές μέθοδοι ανάλυσης χρονοσειρών αποσυνθέτουν την χρονοσειρά σε τέσσερα στοιχεία: την τάση, την κυκλικότητα, την εποχικότητα και τις τυχαίες ή η κανονικές διακυμάνσεις.



Εικόνα 8. Απεικόνιση χαρακτηριστικού παραδείγματος αποσύνθεσης χρονοσειράς

Τάση (T): υποδεικνύει την μακροπρόθεσμη κατεύθυνση μεταβολής του μέσου επιπέδου τιμών της χρονοσειράς. Μια χρονοσειρά χαρακτηρίζεται από αύξουσα , φθίνουσα ή καμία τάση, στις περιπτώσεις που το επίπεδο των τιμών δεν μεταβάλλεται αισθητά. Χρειάζεται προσοχή στην αναγνώριση του στοιχείου της τάσης μιας και απαιτείται μεγάλος αριθμός παρατηρήσεων για να καθοριστεί με ακρίβεια. Είναι πιθανό σε χρονοσειρές που εμφανίζουν κυκλικές μεταβολές, η εξέταση ενός μικρού τμήματος δεδομένων που εμφανίζουν αυξητική τάση να οδηγήσει σε λάθος συμπεράσματα. Έτσι ανάλογα από τα εξεταζόμενα μεγέθη και τα χαρακτηριστικά τους, θα πρέπει να υπάρχουν επαρκή δεδομένα και η τάση να αναζητείται σε ανάλογα διαστήματα δεδομένων.

Κυκλικότητα (C): (ή κυκλικές μεταβολές όπως αναφέρθηκαν παραπάνω) αποτελεί μια περιοδική ή «κυματοειδή» μεταβολή της χρονοσειράς σε μεγάλες γενικά περιόδους χρόνου. Η αιτία για αυτές τις μεταβολές είναι κάποιοι εξωγενείς παράγοντες ενώ απαιτείται πολύ μεγάλος όγκος δεδομένων για να διαπιστωθεί η κυκλική συμπεριφορά ενός μεγέθους. Κυκλική συμπεριφορά παρουσιάζουν οι περισσότεροι οικονομικοί δείκτες όπως το ΑΕΠ μιας χώρας, καθώς και οικονομικά μεγέθη όπως η τιμή του πετρελαίου.

Εποχικότητα (S): ορίζεται ως μια περιοδική διακύμανση στη χρονοσειρά με σταθερή περίοδο που είναι κατά κανόνα μικρότερη του έτους. Τα αίτια της εποχικότητας είναι εύκολα αντιληπτά και επηρεάζουν με τον ίδιο τρόπο τη χρονοσειρά σε κάθε χρονική περίοδο. Η ευκολία στην αναγνώριση και ο σταθερός τρόπος που επηρεάζεται η πορεία της χρονοσειράς καθιστούν εύκολη την απομόνωση της εποχικότητας για την παραγωγή προβλέψεων. Για παράδειγμα, οι αυξημένες πωλήσεις των καταστημάτων λιανικής κατά την εορταστική περίοδο ή η ζήτηση πετρελαίου θέρμανσης τη χειμερινή περίοδο.

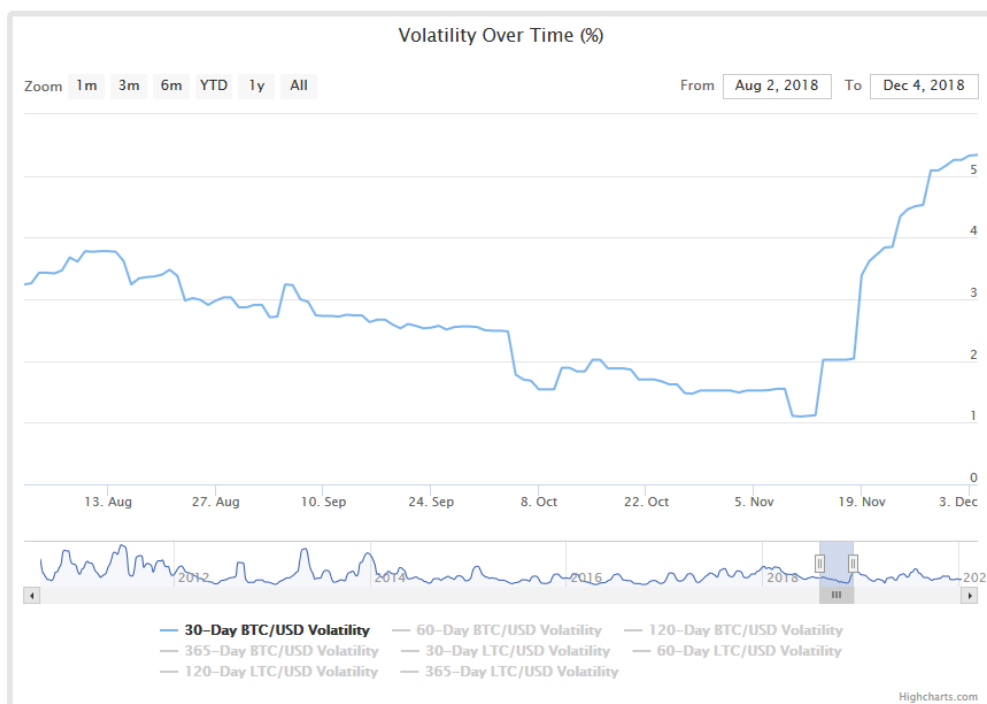
Τυχαιότητα (R): αποτελεί το τελευταίο χαρακτηριστικό της χρονοσειράς αν αφαιρεθούν οι παραπάνω τρεις παράγοντες. Η τυχαιότητα εκφράζει τον στοχαστικό χαρακτήρα της χρονοσειράς πραγματικών δεδομένων. Δεν μπορεί να προβλεφθεί βάσει των προηγούμενων παρατηρήσεων και μπορεί να οφείλεται σε πληθώρα εξωγενών παραγόντων ή ενδογενών απρόβλεπτων συμβάντων. Υπάρχουν περιπτώσεις που η χρονοσειρά εμφανίζει ειδικά γεγονότα με τη μορφή ασυνέχειας ή ακραίων παρατηρήσεων ή ακόμα και αλλαγή επιπέδου. Όπως θα δούμε παρακάτω, η τιμή ενός κρυπτονομίσματος μπορεί να εμφανίσει στιγμιαία μία αρκετά ασυνήθιστη τιμή λόγω ποικίλων παραγόντων. Καθώς η χρονοσειρά που εξετάζεται στην παρούσα εργασία δεν περιέχει έντονη κυκλικότητα ή εποχικότητα, παρά μόνο σποραδικές τάσεις , κρίνεται αναγκαίο να διερευνήσουμε την τυχαιότητα της χρονοσειράς για την επίτευξη

καλύτερων προβλέψεων. Για το λόγο αυτό παρατίθενται παρακάτω 2 διαγράμματα : το 1^ο είναι και

Στα επόμενα διαγράμματα απεικονίζονται η τιμή και η μεταβλητότητα του κρυπτονομίσματος bitcoin για την περίοδο Αύγουστος – Δεκέμβριο 2018, η οποία είναι και η περίοδος που εξετάζεται σε ολόκληρη την παρούσα μελέτη.



Εικόνα 9. Απεικόνιση της τιμής του κρυπτονομίσματος bitcoin για την περίοδο Αύγουστος – Δεκέμβριο 2018 (Πηγή: Highcharts.com)



Εικόνα 10. Απεικόνιση της μεταβλητότητας (volatility) του κρυπτονομίσματος bitcoin για την περίοδο Αύγουστος – Δεκέμβριο 2018 (Πηγή: Highcharts.com)

Συνοψίζοντας όλα τα παραπάνω, η αποσύνθεση της χρονοσειράς στα τέσσερα βασικά στοιχεία αποτελεί ένα πρώτο βήμα στην επεξεργασία και ανάλυση των δεδομένων, ώστε να αποφανθούν κάποια ποιοτικά χαρακτηριστικά της χρονοσειράς και να επιλεγούν μέθοδοι προβλέψεων που ανταποκρίνονται καλύτερα σε αυτά. Έτσι η αναγνώριση και ο διαχωρισμός των τεσσάρων αυτών στοιχείων οδηγεί στην προσαρμογή της χρονοσειράς στα «πραγματικά» δεδομένα και όχι στους εξωγενείς παράγοντες. Στη συνέχεια θα παρουσιαστούν οι βασικές τεχνικές προβλέψεων.

2.3. Βασικές Μέθοδοι Πρόβλεψης

Οι ποσοτικές προβλέψεις μπορούν να χωριστούν σε στατιστικές και κριτικές μεθόδους. Οι στατιστικές μέθοδοι χρησιμοποιούν στατιστικά ή αιτιοκρατικά μοντέλα σε μια σειρά πραγματικών δεδομένων με στόχο την αυτόματη και συστηματική παραγωγή προβλέψεων. Αντίθετα, οι κριτικές μέθοδοι βασίζονται περισσότερο στην γνώση και την εμπειρία των ανθρώπων που παράγουν τις προβλέψεις (REF). Το μεγάλο πλεονέκτημα των κριτικών μεθόδων είναι πως μπορούν να συμπεριληφθούν γεγονότα και ξαφνικές αλλαγές στην διαδικασία πρόβλεψης που τα στατιστικά μοντέλα δεν γίνεται να ενσωματώσουν, ενώ παράλληλα ξεπερνάνε πιθανές ανεπάρκειες σε δεδομένα. Πολλές φορές η κριτική πρόβλεψη εμπειρογνομόνων «διορθώνει» το στατιστικό μοντέλο για καλύτερα αποτελέσματα. Το κύριο μειονέκτημα είναι η έμφυτη τάση των ανθρώπων να είναι αισιόδοξοι ή απαισιόδοξοι. Ο συνδυασμός στατιστικών και κριτικών μεθόδων πρόβλεψης δείχνει να αποτελεί την καλύτερη λύση, κρατώντας τα πλεονεκτήματα της κάθε μεθόδου.

Στα πλαίσια της εν λόγω μελέτης θα χρησιμοποιηθούν σχεδόν αποκλειστικά στατιστικές μέθοδοι προβλέψεων.

Σημαντικό στοιχείο στη επιλογή κατάλληλης μεθόδου στατιστικής πρόβλεψης αποτελεί ο ορίζοντας πρόβλεψης, δηλαδή πόσο μακριά στο μέλλον θέλουμε να παράγουμε πρόβλεψη. Η συνήθης κατηγοριοποίηση είναι η ακόλουθη:

- Βραχυπρόθεσμη πρόβλεψη (inventory forecasting). Αναφέρεται σε σχεδιασμό αποθήκης και η τιμή του ορίζοντα πρόβλεψης είναι σχετικά μικρή (συνήθως μικρότερη από τρεις περιόδους).
- Μεσοπρόθεσμη πρόβλεψη (budget forecasting). Είναι η συνηθέστερη κατηγορία πρόβλεψης και αναφέρεται στον οικονομικό σχεδιασμό της επιχείρησης, όπου ο ορίζοντας πρόβλεψης είναι συνήθως λίγο μεγαλύτερος από ένα οικονομικό έτος (12 με 15 περιόδους για μηνιαίες παρατηρήσεις).

- Μακροπρόθεσμη πρόβλεψη (long term forecasting). Αναφέρεται στο μακροχρόνιο σχεδιασμό των επενδύσεων και της μακροχρόνιας ανάπτυξης. Ο ορίζοντας πρόβλεψης είναι συνήθως μεγαλύτερος από τρία έτη.

Στην εν λόγω εργασία, τα δεδομένα προς επεξεργασία είναι ημερήσια, και συνεπώς, οι έννοιες του ορίζοντα είναι λίγο διαφορετικές. Θα παραχθούν προβλέψεις από 1 μέχρι και 25 χρονικές περιόδους (λίγο πάνω από ένα μήνα δηλαδή) για κάθε χρονοσειρά.

Στη συνέχεια αναλύονται οι κυριότερες τεχνικές προβλέψεων (με “F” συμβολίζεται η πρόβλεψη και με “Y” τα δεδομένα).

2.3.1. Απλοϊκή Μέθοδος (Naïve)

Η «Απλοϊκή ή αφελής» μέθοδος αποτελεί την απλούστερη τεχνική που μπορεί να χρησιμοποιηθεί. Υποδηλώνει δηλαδή ότι τιμή της επόμενης περιόδου θα είναι ίση με την τιμή της προηγούμενης. Η μέθοδος Naïve χρησιμοποιείται κυρίως ως σημείο αναφοράς (benchmark) για σύγκριση με εναλλακτικές πιο σύνθετες μεθόδους. Επίσης, μπορεί να δώσει σχετικά ακριβή αποτελέσματα σε χρονοσειρές που δεν έχουν αποεποχικοποιηθεί, ή δεν έχουν γενικά εποχικότητα καθώς σε αυτές τις χρονοσειρές οι μεταβολές στην τιμή ανά περίοδο είναι μικρές. Η μαθηματική σχέση που περιγράφει τη μέθοδο είναι:

$$F_t = Y_{t-1}$$

2.3.2. Μέθοδος Κινητού Μέσου Όρου

Αρχικά, όπως είναι φανερό πρόκειται για μία μέθοδο πρόβλεψης που βασίζεται στην χρήση του δείκτη μέσου όρου. Ως μέσο όρο ορίζουμε το άθροισμα n τιμών δια του πλήθους τους n. Δηλαδή:

$$\bar{x} = \frac{1}{n} * \sum_{i=1}^n x_i$$

Ο μέσος όρος ενός δείγματος χρησιμοποιείται συχνά σαν αντιπροσωπευτική τιμή αυτού, επομένως είναι εύλογο να ειπωθεί ότι η επόμενη τιμή μιας χρονοσειράς δεν θα απέχει πολύ από το μέσο όρο της. Συνεπώς η πρόβλεψη με τον απλό μέσο όρο είναι:

$$F_{t+1} = \frac{1}{n} * \sum_{i=1}^n Y_i$$

Όσο όμως προχωρά ο χρόνος οι τιμές που λαμβάνουμε υπόψιν για την πρόβλεψη απέχουν πολύ από το επιθυμητό χρονικό ορίζοντα της πρόβλεψης, με αποτέλεσμα να χρησιμοποιούνται παραπάνω πληροφορίες από αυτές που χρειάζονται και οι προβλέψεις να είναι άστοχες. Με γνώμονα αυτό είναι απαραίτητο να υπάρχει η δυνατότητα επιλογής του πλήθους των τιμών που θα λαμβάνονται υπόψιν στην

πρόβλεψη, δημιουργώντας ουσιαστικά ένα παράθυρο(window) με μήκος k , που ορίζεται από τον εκείνον που καλείται να παράξει την πρόβλεψη. Αυτό με μαθηματικούς όρους εκφράζεται ως εξής :

$$F_{t+1} = \frac{1}{k} * \sum_{i=t-k+1}^t Y_i$$

Αξίζει να σημειωθεί ότι η χρήση μεθόδων του μέσου όρου παρά την απλοϊκότητα της χρησιμεύει αρκετές φορές στην εξάλειψη της τάσης και της εποχιακότητας σε περιπτώσεις που θέλουμε να γίνει αποσύνθεση της χρονοσειράς στα βασικά της χαρακτηριστικά

2.3.3. Εκθετική εξομάλυνση

Τα μοντέλα εκθετικής εξομάλυνσης αποτελούν ουσιαστικά επέκταση των μεθόδων μέσου όρου. Στις τελευταίες οι παρατηρήσεις έχουν την ίδια συνεισφορά στην εξαγωγή των προβλέψεων, σε αντίθεση οι μέθοδοι με τις μεθόδους εκθετικής εξομάλυνσης που κάθε παρατήρηση έχει το δικό της βάρος. Ο λόγος για τον οποίο εφαρμόζουμε βάρη μπροστά από κάθε παρατήρηση έγκειται στην ανάγκη να δώσουμε έμφαση σε συγκεκριμένα σημεία στη χρονοσειρά για την εξαγωγή πρόβλεψης. Στις μεθόδους εκθετικής εξομάλυνσης η βαρύτητα δίνεται στις πιο πρόσφατες παρατηρήσεις και μειώνεται εκθετικά όσο οι παρατηρήσεις απομακρύνονται χρονικά από τη στιγμή της πρόβλεψης, γεγονός που έδωσε και το όνομα στις μεθόδους αυτές. Τα μοντέλα εκθετικής εξομάλυνσης χωρίζονται σε 4 κατηγορίες ανάλογα με την ύπαρξη και τη μορφή της τάσης(trens) της προς επεξεργασία χρονοσειράς. Η γενική κατηγοριοποίηση υποδεικνύει τέσσερα μοντέλα τάσης:

- Μοντέλα σταθερού επιπέδου
- Μοντέλα γραμμικής τάσης
- Μοντέλα εκθετικής τάσης
- Μοντέλα φθίνουσας τάσης

Απλή Εκθετική εξομάλυνση (Simple Exponential Smoothing)

Το μοντέλο αυτό χρησιμοποιείται όταν δεν υπάρχει τάση ή υπάρχει σποραδικά σε μεμονωμένα διαστήματα, η χρονοσειρά δηλαδή έχει σταθερό επίπεδο. Στην παρούσα εργασία λόγο της ιδιοσυγκρασίας της χρονοσειράς, η μέθοδος της απλής εκθετικής εξομάλυνσης έδωσε τα καλύτερα αποτελέσματα προβλέψεων.

Οι μαθηματικές σχέσεις που περιγράφουν το μοντέλο είναι:

$$e_t = Y_t - F_t$$

$$S_t = S_{t-1} + a * e_t$$

$$F_{t+1} = S_t$$

Στις παραπάνω εξισώσεις, το e_t δηλώνει το σφάλμα πρόβλεψης, το S το επίπεδο, F την πρόβλεψη και α μια σταθερά εξομάλυνσης που λαμβάνει οποιαδήποτε τιμή στο διάστημα $[0,1]$.

Το αρχικό επίπεδο επιλέγεται με ποικίλους τρόπους, όπως, επιλέγοντας την πρώτη παρατήρηση, το σταθερό επίπεδο από το μοντέλο της απλής γραμμικής παλινδρόμησης ή όπως είναι συνηθέστερο ορίζεται ως αρχικό επίπεδο ο μέσος όρος όλων ή κάποιων αρχικών παρατηρήσεων. Εκτός από τον ορισμό του αρχικού επιπέδου, καταλυτικό ρόλο στην ακρίβεια της πρόβλεψης με απλή εκθετική εξομάλυνση έχει ο καθορισμός του συντελεστή εξομάλυνσης α .

Ο καθορισμός του α εξαρτάται από ορισμένα ποιοτικά χαρακτηριστικά της χρονοσειράς όπως τον θόρυβο (όσο πιο έντονο το φαινόμενο του θορύβου τόσο μικρότερη τιμή παίρνει ο α) και την σταθερότητα του μέσου όρου της χρονοσειράς (μεγάλες μεταβολές αντιμετωπίζονται με μεγαλύτερο α). Για τις ακραία τιμή $\alpha=1$ η πρόβλεψη γίνεται ίδια με την $naive$, ενώ για $\alpha=0$ η πρόβλεψη παραμένει ίδια και ίση με το αρχικό επίπεδο. Υπάρχουν διάφοροι αλγόριθμοι εύρεσης του κατάλληλου α , συνήθως με την εύρεση εκείνου που ελαχιστοποιεί κάποιο δείκτη σφάλματος. Είναι εμφανές πως σε περιπτώσεις που απαιτούνται προβλέψεις ορίζοντα μεγαλύτερου από μια χρονική περίοδο, όλες οι προβλέψεις είναι ίδιες με την τελευταία.

Εκθετική εξομάλυνση με προσαρμογή τάσης (Holt's Method)

Το μοντέλο γραμμικής τάσης (Holt Exponential Smoothing) αποτελεί εξέλιξη της απλής εκθετικής εξομάλυνσης σε συνδυασμό με τη δυνατότητα διαχείρισης δεδομένων που παρουσιάζουν το στοιχείο της τάσης. Οι μαθηματικές σχέσεις που περιγράφουν το μοντέλο είναι :

$$e_t = Y_t - F_t$$

$$S_t = S_{t-1} + T_{t-1} + \alpha * e_t$$

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

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

Το T_t δηλώνει την τάση και το m είναι ο χρονικός ορίζοντας της πρόβλεψης. Η παράμετρος b υποδηλώνει το συντελεστής εξομάλυνσης τάσης και λαμβάνει τιμές στο διάστημα $[0,1]$. Στο μοντέλο αυτό είναι να αναγκαίο να προσδιοριστούν αρχικά τόσο του αρχικού επιπέδο όσο και ο συντελεστής εξομάλυνσης της τάσης. Το αρχικό επίπεδο ορίζεται όπως στην απλή εκθετική εξομάλυνση, ενώ το b αρχικοποιείται ως η διαφορά της ν -στης και της πρώτης παρατήρησης διαιρεμένης με $\nu-1$, ή ως η σταθερά κλίσης από το μοντέλο της απλής γραμμικής παλινδρόμησης. Το αρχικό επίπεδο και η

αρχική τάση πρέπει να καθορίζονται με προσοχή καθώς επηρεάζουν αρκετά την τελική πρόβλεψη.

Η μεγάλη διαφορά του μοντέλου αυτού από τη μέθοδο SES είναι η παραγωγή προβλέψεων με χρονικό ορίζοντα μεγαλύτερο της μονάδας. Λόγω της θεώρησης πως τα δεδομένα έχουν μια σταθερά ανοδική τάση, οι προβλέψεις για ορίζοντα μεγαλύτερο της μονάδας προκύπτουν με τη χρήση των τελευταίων διαθέσιμων τιμών για το επίπεδο και την τάση και αύξηση του δείκτη m .

Μοντέλα μη γραμμικής τάσης

Το μοντέλο μη γραμμικής τάσης είναι ουσιαστικά μία τροποποίηση του προηγούμενου μοντέλου γραμμικής τάσης. Η ανάγκη τροποποίησης του μοντέλου Holt εκπηγάζει από το γεγονός ότι εμφανίζει συχνά μία τάση για υπερεκτίμηση, με άλλα λόγια οι προβλέψεις που παράγει χαρακτηρίζονται από θετική προκατάληψη, ιδιαίτερα σε περίπτωση μεσοπρόθεσμων ή μακροπρόθεσμων προβλέψεων. Η επέκταση της μεθόδου επιτυγχάνεται με τη προσθήκη μίας επιπλέον παραμέτρου η οποία περιορίζει ουσιαστικά το ρυθμό αύξησης των τιμών. Η παράμετρος αυτή καλείται παράμετρος διόρθωσης της τάσης (trend-modification parameter) και συμβολίζεται ϕ . Οι εξισώσεις που περιγράφουν το συγκεκριμένο μοντέλο με τον τρόπο που είναι οι ακόλουθες (Gardner Jr., E.S. and McKenzie, E. (1985) Forecasting Trends in Time Series):

$$e_t = Y_t - F_t$$

$$S_t = S_{t-1} + \alpha * T_{t-1} + a * e_t$$

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

$$F_{t+m} = S_t + \sum_{i=1}^m \phi^i * e_t$$

Πέραν των γνωστών μεταβλητών που περιεγράφηκαν στα παραπάνω δύο μοντέλα, η μέθοδος αυτή περιλαμβάνει ακόμη το ϕ , το οποίο εκφράζει διόρθωσης της τάσης.

Όπως είναι εμφανές ο υπολογισμός της τάσης σε αυτό το μοντέλο είναι μη γραμμικός. Σε αντίθεση με τις παραμέτρους α και β , η παράμετρος ϕ μπορεί να λάβει οποιαδήποτε μεγαλύτερη του μηδενός τιμή, παρόλα αυτά είναι αναγκαίο να περιοριστεί με τη χρήση άνω και κάτω ορίων για ορισμένες χαρακτηριστικές τιμές. Επομένως, μπορούμε να χαρακτηρίσουμε το μοντέλο μη γραμμικής τάσης ανάλογα με την τιμή που παίρνει κάθε φορά η παράμετρος ϕ , ως εξής :

- Αν $\varphi = 0$, τότε προκύπτει το μοντέλο της απλής εκθετικής εξομάλυνσης (SES), διότι η τάση δεν έχει πλέον καμία επίδραση στον κ καθορισμό των στατιστικών σημειακών προβλέψεων.
- Αν $0 < \varphi < 1$, τότε προκύπτει το μοντέλο φθίνουσας τάσης (damped exponential smoothing), το οποίο θεωρείται καταλληλότερο όταν μας ενδιαφέρει η παραγωγή μεσοπρόθεσμων προβλέψεων, καθώς χαρακτηρίζεται από έλλειψη τάσης για υπεραισιοδοξία.
- Αν $\varphi = 1$, τότε προκύπτει το μοντέλο της γραμμικής τάσης, αφού στην τελευταία εξίσωση η οποία παράγει και την πρόβλεψη, τη θέση του αθροίσματος παίρνει το γινόμενο της μεταβλητής χρονικού ορίζοντα m και τα προηγούμενης τάσης T_t .
- Αν $\varphi > 1$, τότε προκύπτει το μοντέλο της εκθετικής τάσης, το οποίο χαρακτηρίζεται από μεγάλη θετική προκατάληψη και χρησιμοποιείται μόνο σε περιπτώσεις όπου το ζητούμενο είναι η πρόβλεψη τιμής στην αρχή του κύκλου ζωής ενός προϊόντος ή μιας υπηρεσίας.

Το μεγάλο μειονέκτημα του συγκεκριμένου μοντέλου, είναι η παραγωγή ιδιαίτερα προκατειλημμένων προβλέψεων, κάτι το οποίο ενδέχεται να προκύψει από λανθασμένη επιλογή του συντελεστή φ . Γι' αυτό το λόγο, συνίσταται ο περιορισμός της παραμέτρου φ στο διάστημα $[0,1]$. Αξίζει να σημειωθεί πως η ακρίβεια των προβλέψεων που παράγονται από το μοντέλο μη γραμμικής τάσης είναι σημαντικά μεγαλύτερη σε σχέση με αυτές του μοντέλου γραμμικής τάσης.

Γενικά, όσο μεγαλύτερος είναι ο απαιτούμενος ορίζοντας της πρόβλεψης, τόσο μεγαλύτερο πλεονέκτημα παρουσιάζει το συγκεκριμένο μοντέλο έναντι των υπόλοιπων μοντέλων εκθετικής εξομάλυνσης. Όσον αφορά τη βέλτιστη επιλογή των παραμέτρων α , β και φ , αυτή επιτυγχάνεται με μη γραμμική (δυναμική) αναζήτηση, με κριτήριο πάντα την ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος (MSE), ενώ η επιλογή αρχικού επιπέδου και αρχικής τάσης πραγματοποιείται με εφαρμογή της απλής γραμμικής παλινδρόμησης, όπως ακριβώς και στις άλλες μεθόδους εκθετικής εξομάλυνσης.

2.3.4. Μοντέλα αυτοσυσχέτισης (Arima model)

Τα Ολοκληρωμένα Αυτοπαλίνδρομα μοντέλα Κινητού Μέσου όρου (Autoregressive Integrated Moving Average) είναι στοχαστικά μαθηματικά μοντέλα τα οποία μπορούν να περιγράψουν ικανοποιητικά τις διάφορες συνιστώσες της χρονοσειράς. Τα μοντέλα ARIMA έχουν πολύ ικανοποιητικά αποτελέσματα όσον αφορά την προβλεπτική τους ικανότητα σε βραχυπρόθεσμες προβλέψεις όπως περιγράφεται παρακάτω. Στη βιβλιογραφία αναφέρονται συχνά και σαν μεθοδολογία BOX-JENKINS, χάρη στην εκτεταμένη έρευνα και ανάλυση των παραπάνω στο βιβλίο

τους Time series analysis: Forecasting and control ,George E. P. Box, Gwilym M. Jenkins, 1970.

Πιο συγκεκριμένα, κάθε μοντέλο ARIMA εκφράζεται ως γραμμικός συνδυασμός των παραγόντων αυτοπαλινδρόμησης (AutoRegression) και κινητού μέσου όρου (Moving Average), με στόχο να προσδιοριστεί εκείνος που παράγει τις καλύτερες προβλέψεις. Έτσι, αν το μοντέλο περιλαμβάνει αποκλειστικά παράγοντες αυτοπαλινδρόμησης αναφέρεται ως AR(p), αν περιλαμβάνει αποκλειστικά παράγοντες κινητών μέσω όρων ως MA(q), και αν περιλαμβάνει και τους δύο ως ARMA(p,q), όπου τα p και q δηλώνουν την τάξη του μοντέλου ανά παράγοντα. Ο παράγοντας I(d) αναφέρεται στη διαφόριση της χρονοσειράς πριν την εφαρμογή ενός μοντέλου ARMA(p,q) και έχει ως στόχο την αφαίρεση της τάσης από τα δεδομένα, δηλαδή να κάνει το μοντέλο στάσιμο(stationary) . Σε περίπτωση που η χρονοσειρά είναι εποχιακή, τα μοντέλα ARIMA μπορούν να επεκταθούν κατάλληλα προκειμένου να προσομοιώσουν και την εποχιακή συμπεριφορά των δεδομένων. Σε αυτή την περίπτωση η έκφρασή τους έχει τη μορφή ARIMA(p,d,q)(P,D,Q), όπου τα P, D και Q αναφέρονται αντίστοιχα στην τάξη των εποχιακών παραγόντων ARIMA.

Στην παρούσα διπλωματική εργασία παρουσιάζεται συνοπτικά η μεθοδολογία πρόβλεψης του ολοκληρωμένου αυτοπαλινδρομικού μοντέλου κινητού μέσου όρου.

1.6.3.1. Μοντέλα αυτοπαλινδρόμησης-AR(p)

Τα μοντέλα αυτοπαλινδρόμησης θεωρούν γραμμικές σχέσεις ανάμεσα στις παρατηρήσεις της ίδιας της χρονοσειράς και τις αξιοποιούν για την περιγραφή της και την παραγωγή προβλέψεων. Ένα τέτοιο μοντέλο αναφέρεται ως μοντέλο αυτοπαλινδρόμησης AR(p), όπου p η τάξη του, δηλαδή το πλήθος των παρελθοντικών παρατηρήσεων που αξιοποιούνται για την παραγωγή της πρόβλεψης. Αλγεβρικά αυτό αναπαρίσταται ως εξής:

$$y_t = c + \varphi_1 * y_{t-1} + \dots + \varphi_p * y_{t-p} \quad (1),$$

όπου φ_i οι συντελεστές αυτοσυσχέτισης του μοντέλου για υστέρηση i και c μία σταθερά. Υπολογίζεται ως γραμμικός συνδυασμός αυτών προσαυξανόντάς την προαιρετικά κατά μία σταθερά c .Χρησιμοποιώντας τον τελεστή ολίσθησης, ένα μοντέλο AR μπορεί να γραφτεί και ως:

$$(1 - \varphi_1 * B - \varphi_2 * B^2 - \dots - \varphi_p * B^p) * \bar{y}_t = 0,$$

όπου $\bar{y}_t = y_t - \mu$. Η χρονοσειρά \bar{y}_t έχει τις ίδιες στατιστικές ιδιότητες με την αρχική χρονοσειρά με μηδενική μέση τιμή. Η χρήση της γίνεται προκειμένου να τονιστούν οι στοχαστικές συνιστώσες της χρονοσειράς. Αν αναπτυχθεί η παραπάνω σχέση έχουμε:

$$y_t = \mu(1 - \varphi_1 - \varphi_2 - \dots - \varphi_p) + \varphi_1 * y_{t-1} + \varphi_2 * y_{t-2} + \dots + \varphi_p * y_{t-p} \quad (2)$$

Από τις σχέσεις (1) και (2) οδηγούμαστε στο συμπέρασμα ότι για τη σταθερά c σε ένα μοντέλο AR(p) ισχύει

$$c = \mu(1 - \varphi_1 - \varphi_2 - \dots - \varphi_p)$$

1.6.3.2. Μοντέλα κινητού μέσου όρου-MA(q)

Τα μοντέλα κινητού μέσου όρου περιγράφησαν παραπάνω, παρόλα αυτά θα αναφερθούν ξανά με γνώμονα την ευκολότερη κατανόηση του ολοκληρωμένου μοντέλου και τους συμβολισμούς που αυτό περιλαμβάνει. Επομένως η μέθοδος των κινητών μέσων όρων ο γράφεται αλγεβρικά ως εξής:

$$\bar{y}_t = c + \theta_1 * e_{t-1} + \theta_2 * e_{t-2} + \dots + \theta_q * e_{t-q} \quad (3),$$

όπου θ_i οι συντελεστές μερικής αυτοσυσχέτισης του μοντέλου MA για υστέρηση i .

Χρησιμοποιώντας τον τελεστή ολίσθησης, ένα μοντέλο MA μπορεί να γραφτεί και ως:

$$\bar{y}_t = (\theta_1 * B + \theta_2 * B^2 + \dots + \theta_q * B^q) * e_t$$

Αν αναπτύξουμε τώρα την παραπάνω σχέση έχουμε

$$\bar{y}_t = \mu + \theta_1 * e_{t-1} + \theta_2 * e_{t-2} + \dots + \theta_q * e_{t-q} \quad (4)$$

Από τις σχέσεις (3) και (4) οδηγούμαστε στο συμπέρασμα ότι για τη σταθερά c σε ένα μοντέλο MA(q) ισχύει $c = \mu$.

1.6.3.3. Ολοκληρωμένα αυτοπαλινδρομικά μοντέλα κινητού μέσου όρου -ARIMA(p,d,q)

Το συνολικό μοντέλο αναπαρίσταται με τη χρήση του τελεστή ολίσθησης B ως εξής:

$$(1 - \varphi_1 * B - \varphi_2 * B^2 - \dots - \varphi_p * B^p) * (1 - B)^n * (1 - B^m)^N * y_t = c + (\theta_1 * B + \theta_2 * B^2 + \dots + \theta_q * B^q) e_t$$

Ο πρώτος όρος του πρώτου μέλους της εξίσωσης αναπαριστά το μοντέλο AR(p), ο δεύτερος την διαφόριση I(d), ενώ ο όρος στο δεύτερο μέλος της εξίσωσης το μοντέλο MA(q).

Θέλοντας να προσδιορίσουμε την σταθερά c , εξισώνουμε τις δύο παρακάτω σχέσεις: $(1 - \varphi_1 * B - \varphi_2 * B^2 - \dots - \varphi_p * B^p) * (1 - B)^n * (1 - B^m)^N * \bar{y}_t = c + (\theta_1 * B + \theta_2 * B^2 + \dots + \theta_q * B^q) e_t$

$$\bar{y}'_t = c + \varphi_1 * y'_{t-1} + \varphi_2 * y'_{t-2} + \dots + \varphi_p * y'_{t-p} + \theta_1 * e_{t-1} + \theta_2 * e_{t-2} + \dots + \theta_q * e_{t-q}$$

, που \bar{y}'_t το προϊόν της εποχιακής και μη διαφόρισης της χρονοσειράς. Αναπτύσσοντας τώρα την πρώτη σχέση κατά τα γνωστά προκύπτει ότι:

$$\begin{aligned}
& y_t * (1 - \varphi_1 - \varphi_2 - \dots - \varphi_p)(1 - B)^n(1 - B^m)^N \\
& - \mu(1 - \varphi_1 - \varphi_2 - \dots - \varphi_p)(1 - B)^n(1 - B^m)^N \\
& = \theta_1 B^e t + \theta_2 B^2 e t + \dots + \theta_q B^q e t - q
\end{aligned}$$

Η σταθερά c ισούται προφανώς με τον αντίθετο του δεύτερου όρου του πρώτου μέλους της παραπάνω εξίσωσης. Έτσι προκύπτει ότι για $n=N=0$ για την σταθερά c θα ισχύει:

$$c = \mu(1 - \varphi_1 - \varphi_2 - \dots - \varphi_p)$$

Σε οποιαδήποτε άλλη περίπτωση η σταθερά ισούται με μηδέν. Αυτό ήταν αναμενόμενο δεδομένου ότι ιδανικά ο μέσος όρος μίας διαφορισμένης και άρα στάσιμης χρονοσειράς ισούται με μηδέν. Πρακτικά η χρονοσειρά που προκύπτει από τη διαφορίση δεν είναι ποτέ απολύτως στάσιμη γύρω από το μηδέν. Έτσι, συνηθίζεται να προστίθεται σταθερά ακόμα και σε χρονοσειρές που έχουν υποστεί διαφορίση. Συνοπτικά αναφέρονται τα εξής:

- Η μη διαφορίση ($d=0$) είναι δείγμα ύπαρξης σταθερότητας στην αρχική χρονοσειρά. Ωστόσο αυτή η σταθερότητα ενδέχεται να μην έχει για κέντρο της το μηδέν. Η εισαγωγή μίας σταθεράς μπορεί να βοηθήσει στον καλύτερο προσδιορισμό του πραγματικού επιπέδου της χρονοσειράς και συνεπώς να βελτιώσει προβλεπτικά την επίδοσή του.
- Η διαφορίση πρώτης τάξης ($d=1$) είναι δείγμα ύπαρξης τάσης στην αρχική χρονοσειρά. Αυτή εξαλείφεται κατά τη διαφορίση, οπότε και οι προβλέψεις που παράγονται στη συνέχεια κινούνται γύρω από ένα σταθερό επίπεδο. Η εισαγωγή σταθεράς μπορεί να βοηθήσει στην επανένταξη της τάσης που αφαιρέθηκε κατά τη διαφορίση στις τελικές προβλέψεις και συνεπώς να βελτιώσει προβλεπτικά την επίδοση του μοντέλου.
- Η διαφορίση δεύτερης τάξης ($d=2$) συνεπάγεται ύπαρξη χρονικά μεταβαλλόμενης τάσης (τάση μέσα στην τάση) στην αρχική χρονοσειρά. Έτσι, η εισαγωγή σταθεράς σε αυτήν την περίπτωση θεωρείται άστοχη επιλογή, εκτός και αν υπάρχει σοβαρή ένδειξη ύπαρξης εκθετικού μοτίβου ανάπτυξης.

Παρακάτω δίνονται οι περιορισμοί που πρέπει να πληρούνται από τα πιο διαδεδομένα μοντέλα ARMA.

- AR(1): $-1 < \varphi_1 < 1$
- AR(2): $-1 < \varphi_2 < 1$ και $\varphi_1 + \varphi_2 < 1$ και $\varphi_2 - \varphi_1 < 1$
- MA(1): $-1 < \theta_1 < 1$
- MA(2): $-1 < \theta_2 < 1$ και $\theta_1 + \theta_2 > -1$ και $\theta_1 - \theta_2 < 1$.

2.4. Ακρίβεια προβλέψεων (Σφάλματα)

Από την στατιστική προέκυψε η ανάγκη κατηγοριοποίησης των γραφικών παραστάσεων τυχαίων μεταβλητών, καθώς και η ανάγκη για εύρεση των χαρακτηριστικών τους. Προέκυψαν, λοιπόν, στατιστικοί δείκτες που εφαρμόζονται πέρα από τη στατιστική σε πολλά άλλα πεδία επιστημών, όπως αυτό της λήψης αποφάσεων (decision making) και των προβλέψεων. Στις προβλέψεις ειδικότερα, αφορά σε στατιστική ανάλυση χρονοσειράς που βοηθούν στη συλλογή πληροφορίας και στην επιλογή κατάλληλης στρατηγικής επίλυσης ενός προβλήματος απόφασης και πρόβλεψης.

Παρακάτω παρουσιάζονται οι πιο βασικοί στατιστικοί δείκτες και περιγράφονται οι βασικές ιδιότητές τους.

- Μέση τιμή (Mean Value): Η μέση τιμή, όπως έχει αναφερθεί, είναι ένα μέτρο που εκφράζει την αντιπροσωπευτική τιμή μιας κατανομής τιμών. Περιγράφετε ως:

$$\bar{Y} = \frac{1}{n} * \sum_{i=1}^n Y_i$$

- Μέγιστη και Ελάχιστη τιμή (Maximum, Minimum): Λέγονται και ακρότατα μιας συνάρτησης, εκφράζουν τη μεγαλύτερη και τη μικρότερη τιμή σε μια κατανομή. Η γνώση τους βοηθά στο να φράξει κάποιος το εύρος τιμών που μπορεί να κινηθεί το υπό εξέταση μέγεθος.
- Τυπική απόκλιση (Standard Deviation): Η μέση τιμή δεν μπορεί να περιγράψει την απόκλιση των τιμών από την κεντρική τιμή, αυτό επιτυγχάνεται μερικώς με την τυπική απόκλιση.

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n}}$$

- Διακύμανση (Variance): Τετράγωνο της τυπικής απόκλισης. Ο υπολογισμός είναι πιο εύκολος, όμως δεν διατηρεί τις μονάδες μέτρησης.

$$Var = \sigma^2$$

- Συνδιακύμανση (Covariance): Εκφράζει την αλληλεξάρτηση μεταξύ δύο μεταβλητών, αν η μεταβολή στη μία προκαλεί μεταβολή στην άλλη, καθώς και το πρόσημο της μεταβολής.

$$COV(X, Y) = \frac{1}{n} * \sum_{i=1}^n [(X_i - \bar{X})(Y_i - \bar{Y})]$$

Θετική συνδιακύμανση σημαίνει ότι τα μεγέθη μεταβάλλονται ανάλογα, αρνητική ότι τα μεγέθη είναι αντιστρόφως ανάλογα και μηδενική ότι τα μεγέθη είναι ασυσχέτιστα.

- Συντελεστής γραμμικής συσχέτισης (Linear Correlation Coefficient): Εκφράζει την συγκέντρωση των σημείων ενός διαγράμματος διασποράς γύρω από την ευθεία παλινδρόμησης, δίνοντας έτσι ένα μέτρο της γραμμικής συσχέτισης μεταξύ των δύο μεταβλητών.

$$r_{XY} = \frac{\sum_{i=1}^n [(X_i - \bar{X})(Y_i - \bar{Y})]}{\sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2} * \sqrt{\sum_{i=1}^n (X_i - \bar{X})^2}}$$

Για $r_{XY} = \pm 1$ έχουμε τέλεια γραμμική συσχέτιση.

Για $-0,3 < r_{XY} < 0,3$ δεν υπάρχει γραμμική συσχέτιση.

- Συντελεστής αυτοσυσχέτισης (Autocorrelation Coefficient): Εκφράζει τη συσχέτιση μεταξύ παρατηρήσεων της ίδιας μεταβλητής με χρονική υστέρηση k περιόδων. Χρησιμοποιείται στον εντοπισμό εποχιακής συμπεριφοράς.

$$ACF_K = \frac{\sum_{i=k+1}^n [(X_i - \bar{X})(X_{i-k} - \bar{X})]}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

Για $ACF_K = 0$ έχουμε μηδενική συσχέτιση των παρατηρήσεων χρονικής υστέρησης k .

Για $ACF_K = 1$ έχουμε μεγάλη συσχέτιση των παρατηρήσεων χρονικής υστέρησης k .

- Συντελεστής Μεταβλητότητας (Coefficient of Variation): Κανονικοποιημένο μέτρο διασποράς των παρατηρήσεων

$$CV = \sigma/\mu$$

Σε αντίθεση με τις παραπάνω μετρικές της βασικής στατιστικής ανάλυσης που ενεργούν πάνω στην πραγματική σειρά των παρατηρήσεων, ορίζουμε νέες μετρικές, χρήσιμες για το αντικείμενο προβλέψεων, που χρησιμοποιούν την πραγματική σειρά

των παρατηρήσεων αλλά και μία σειρά τιμών πρόβλεψης επί της πραγματικής χρονοσειράς. Η εφαρμογή μίας μεθόδου πρόβλεψης έχει ως αποτέλεσμα την παραγωγή τιμών h τιμών, όπου h ο ορίζοντας πρόβλεψης. Γενικότερα, το σύνολο των τιμών πρόβλεψης αποτελείται από δύο μέρη, τις πρώτες n παρατηρήσεις για τις οποίες έχουμε και πρόβλεψη και πραγματική τιμή και ονομάζεται προσαρμογή μοντέλου πρόβλεψης (fitted model), ενώ οι υπόλοιπες είναι απλά μελλοντικές προβλέψεις. Μέχρι και τη n -οστή παρατήρηση το δείγμα λέγεται “in-sample” και μπορούμε να υπολογίσουμε το σφάλμα σε αυτό, ενώ από $n + 1$ έως $n + h$ μιλάμε για “out-of-sample” δηλαδή την περιοχή όπου δεν έχουμε πραγματικές τιμές ακόμη, μόνο προβλέψεις.

Οι δείκτες που παρουσιάζονται στη συνέχεια αναφέρονται στο “in-sample-error” μόνο και είναι οι εξής:

- Μέσο σφάλμα (Mean Error): $ME = 1/n * \sum_{i=1}^n (Y_i - \overline{F_i})$

- Μέσο απόλυτο σφάλμα (Mean Absolute Error):

$$MAE = 1/n * \sum_{i=1}^n |Y_i - \overline{F_i}|$$

- Μέσο τετραγωνικό σφάλμα (Mean Squared Error):

$$MSE = \frac{1}{n} * \sum_{i=1}^n (Y_i - \overline{F_i})^2$$

- Ρίζα Μέσου τετραγωνικού σφάλματος (Root Mean Squared Error):

$$\sqrt{MSE} = \sqrt{\frac{1}{n} * \sum_{i=1}^n (Y_i - \overline{F_i})^2}$$

- Μέσο απόλυτο ποσοστιαίο σφάλμα (Mean Absolute Percentage Error):

$$MAPE = \frac{1}{n} * \sum_{i=1}^n \left| \frac{Y_i - F_i}{Y_i} \right| * 100\%$$

- Συμμετρικό μέσο απόλυτο ποσοστιαίο σφάλμα (Symmetric Mean Absolute

$$Percentage Error): sMAPE = \frac{1}{n} * \sum_{i=1}^n \left| \frac{Y_i - F_i}{\frac{(Y_i + F_i)}{2}} \right| * 100\%$$

Εν κατακλείδι, σχετικά με τα παραπάνω σφάλματα, παρατηρείται πως τα: Mean Error (ME), Mean Absolute Error (MAE), Mean Squared Error (MSE) και Root Mean Squared Error (RMSE) έχουν ως πυρήνα τις διαφορές των προβλέψεων από τις αντίστοιχες πραγματικές τιμές. Τα Mean Absolute Percentage Error (MAPE) και Symmetric Mean Absolute Percentage (sMAPE) εκφράζουν σφάλματα σε ποσοστιαία μορφή, ενώ το Mean Absolute Scaled Error (MASE) αποτελεί σχετικό μέτρο που χρησιμοποιεί το σφάλμα μιας μεθόδου βάσης και το συγκρίνει με το σφάλμα της υπό εξέταση μεθόδου. Στην παρούσα διπλωματική εργασία χρησιμοποιήθηκε το Mean Absolute Error (MAE) για την αξιολόγηση των προβλέψεων. Στη συνέχεια θα

παρουσιαστεί θεωρητικά το υπόβαθρο της στρατηγικής που ακολουθήθηκε για την παραγωγή των προβλέψεων.

3. Μέθοδοι Συνάθροισης Δεδομένων (Aggregation)

Η συνάθροιση δεδομένων (data aggregation) είναι η διαδικασία κατά την οποία έχουμε τη δυνατότητα να αναδείξουμε τα ποιοτικά χαρακτηριστικά μιας χρονοσειράς σε διαφορετικά επίπεδα χρόνου, ή διαφορετικές συχνότητες.

Τα τρία βασικότερα είδη συνάθροισης δεδομένων είναι τα εξής:

- Ιεραρχική συνάθροιση (cross-sectional aggregation)
- Χρονική συνάθροιση (temporal aggregation or flow scheme)

Βασική παράμετρο στην υλοποίηση της διαδικασίας αυτής αποτελεί το επίπεδο συνάθροισης (aggregation level) που ορίζεται ανά περίπτωση. Επιλέγεται συνήθως τέτοιο επίπεδο συνάθροισης ώστε να αναδεικνύονται τα ποιοτικά χαρακτηριστικά και οι ιδιαιτερότητες της χρονοσειράς. Στην περίπτωση που το επίπεδο συνάθροισης 'κ' δεν είναι ακέραιο πολλαπλάσιο του πλήθους n των παρατηρήσεων, τότε είθισται να αφαιρούμε 'κ' πλήθος παρατηρήσεων και να αξιοποιούμε τις υπόλοιπες 'n-κ' παρατηρήσεις για την παραγωγή των προβλέψεων.

Οι παρατηρήσεις που αφαιρούνται είναι οι αρχαιότερες χρονικά, εκείνες δηλαδή που το βάρος της επιρροής τους στις επικείμενες προβλέψεις αναμένεται να είναι μικρό. Σύμφωνα με μελέτες έχει προκύψει ότι δεν υπάρχει κάποιος συγκεκριμένος και βέλτιστος τρόπος ορισμού του κατάλληλου επιπέδου συνάθροισης αλλά, αντιθέτως, είναι αναγκαίο κάθε φορά να αναλύονται οι ποικίλοι παράγοντες που χαρακτηρίζουν και επηρεάζουν τη χρονοσειρά. Πιο συγκεκριμένα, παράγοντες όπως το πλήθος των διαθέσιμων ενδογενών χαρακτηριστικών καθώς και η ανομοιογένεια που παρουσιάζουν πρέπει να ληφθούν σοβαρά υπόψη στην επιλογή του επιπέδου.

Οι εν λόγω παράγοντες αποδεικνύουν ότι όσο μεγαλύτερη είναι η ανομοιογένεια της χρονοσειράς, τόσο πιο δύσκολο είναι να εφαρμοστεί η μέθοδος της συνάθροισης. Στη συνέχεια, παρουσιάζεται μια μέθοδος η οποία συνδυάζει τα δεδομένα με βάση τον βαθμό των ομοιοτήτων που παρουσιάζουν και καταλήγουν σε θεμιτά αποτελέσματα για τα δεδομένα της συγκεκριμένης έρευνας, με προοπτικές γενίκευσης σε αντίστοιχων χαρακτηριστικών χρονοσειρές.

3.1. Ιεραρχική Συνάθροιση

Ως ιεραρχική συνάθροιση (cross-sectional aggregation) ορίζεται η διαδικασία κατά την οποία πραγματοποιούνται ξεχωριστά οι προβλέψεις για κάθε χρονοσειρά και στη συνέχεια συνδυάζονται για να την παραγωγή κάποιων νέων προβλέψεων με το χαρακτηριστικό ότι αυτές είναι συμφλιωμένες, δηλαδή το άθροισμά τους ανά επίπεδο είναι ίδιο (Improving forecasting via multiple temporal aggregation, N. Kourentzes, F. Petropoulos, 2014). Χαρακτηριστικό παράδειγμα αυτής της τεχνικής είναι η παραγωγή πρόβλεψης ζήτησης ή τιμής για κάποια παρεμφερή προϊόντα ξεχωριστά και στην συνέχεια ο συνδυασμός τους για την δημιουργία μιας συνολικής συμφλιωμένης

εκτίμησης για το χαρακτηριστικό του δεδομένου προϊόντος. Η μέθοδος της ιεραρχικής συνάθροισης μπορεί να εφαρμοστεί σε περιπτώσεις όπου χρειάζεται ιεραρχική πρόβλεψη (Hierarchical Forecasting), η οποία πραγματοποιείται συνήθως μέσω των top-down και bottom-up (Hierarchical forecasts for Australian domestic tourism, G. Athanasopoulos and R.A. Ahmed, 2009).

Στη συγκεκριμένη μελέτη των G. Athanasopoulos και R.A. Ahmed αναλύθηκε η ιεραρχική φύση των δεδομένων της τουριστικής ζήτησης και παρήχθησαν βραχυπρόθεσμες προβλέψεις για τον εγχώριο τουρισμό στην Αυστραλία. Τόσα τα δεδομένα που δημιουργούν τις χρονοσειρές όσο και οι προβλέψεις τους είναι οργανωμένα σε ιεράρχηση με γνώμονα τη γεωγραφική περιοχή και τον σκοπό του ταξιδιού.

Η ανάλυση πραγματοποιήθηκε με τη χρήση πέντε μεθόδων ιεραρχικής πρόβλεψης. Αρχικά, δημιουργήθηκαν προβλέψεις αξιοποιώντας τις μεθόδους 'top-down' και 'bottom-up', καθώς και μιας νέας μορφής 'top-down', όπου οι 'top-level' προβλέψεις χωρίζονται σύμφωνα με τις προβλεπόμενες αναλογίες των 'lower-level' σειρών. Τέλος, ερευνήθηκε ένας συνδυασμός των παραπάνω. Όπως προέκυψε από τις επιδόσεις των προβλέψεων, οι δύο τελευταίες τεχνικές ήταν οι πιο αποδοτικές για τις ιεραρχίες τουρισμού που χρησιμοποιήθηκαν. Το κυριότερο πλεονέκτημα των δύο νέων προσεγγίσεων είναι η αποτελεσματικότερη σύλληψη των διαφόρων χαρακτηριστικών των σειρών. Για την αξιολόγηση των επιδόσεων των νέων τεχνικών έγινε σύγκριση των αποτελεσμάτων χρησιμοποιώντας δεδομένα εκτός του δείγματος (out of sample data) και το τεστ DieHold-Mariano. Διαπιστώθηκε ότι τα αποτελέσματα των δύο αυτών τεχνικών είχαν σχετικά την ίδια ακρίβεια πρόβλεψης.

Σε ότι αφορά τη σύγκριση μεταξύ 'top-down' και 'bottom-up' οι απόψεις διίστανται. Αυτό προκύπτει από το γεγονός ότι, για παράδειγμα, σε δεδομένα διακοπτόμενης ζήτησης η 'bottom-up' μέθοδος καταλήγει σε πιο ακριβείς προβλέψεις, ενώ στην αναζήτηση συγκεκριμένων μοτίβων στην ζήτηση μεμονωμένων αγαθών η 'top-down' μέθοδος φαίνεται να υπερέχει αισθητά.

Το παραπάνω ερώτημα περί υπεροχής της μίας μεθόδου έναντι της άλλης διαφωτίστηκε όταν οι R.J. Hyndman και R.A. Ahmed (2011) με τη δημοσίευση του «Optimal combination forecasts for hierarchical time series» πρότειναν μια εναλλακτική τεχνική προβλέψεων για ιεραρχικές χρονοσειρές, πιο αποτελεσματική από τις 'top-down' και 'bottom-up'. Η εν λόγω μέθοδος ονομάζεται 'optimal' και βασίζεται σε ανεξάρτητες προβλέψεις που αναφέρονται σε όλα τα επίπεδα ιεραρχίας οι οποίες στην συνέχεια εισάγονται σε ένα μοντέλο παλινδρόμησης για να συνδυαστούν βέλτιστα και να συγχωνευτούν δίνοντας συμφιλιωμένες προβλέψεις. Η μέθοδος αυτή παρέχει λοιπόν σημειακές προβλέψεις οι οποίες συμφιλιώνονται σε όλα τα επίπεδα ιεραρχίας, επιτρέποντας έτσι τους συσχετισμούς και τις αλληλεπιδράσεις μεταξύ των σειρών σε κάθε επίπεδο της ιεραρχίας. Ως εκ τούτου, η μέθοδος είναι αρκετά πιο ευέλικτη σε σχέση με τις 'top-down' και 'bottom-up' μεθόδους. Για την

πραγματοποίηση της έρευνας χρησιμοποιήθηκαν οι χρονοσειρές από την τουριστική ζήτηση στην Αυστραλία τα οποία είχαν ταξινομηθεί ανάλογα με την περιοχή και τον σκοπό του ταξιδιού. Αξίζει να σημειωθεί ότι αυτή η μέθοδος δεν δεσμεύει την επιλογή του μοντέλου πρόβλεψης για την πραγματοποίηση των αρχικών προβλέψεων.

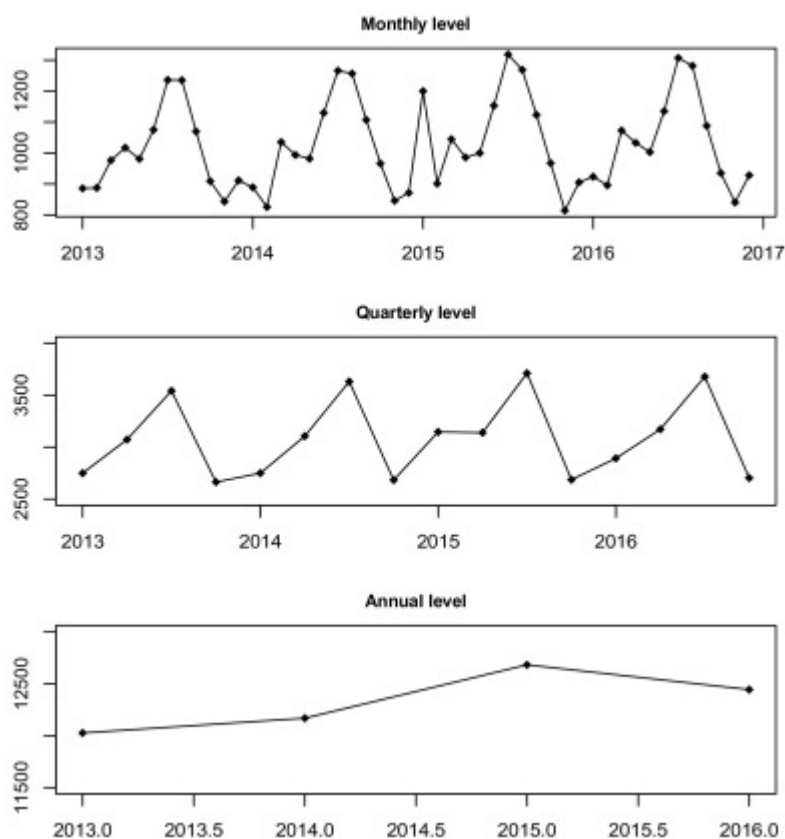
3.2. Χρονική Συνάθροιση

Η δεύτερη μορφή συνάθροισης είναι η χρονική συνάθροιση δεδομένων (temporal aggregation or flow scheme), όπου μετατρέπουμε μια υψηλής συχνότητας χρονοσειρά σε μια χαμηλής συχνότητας. Η μέθοδος αυτή χωρίζεται σε δυο κατηγορίες, την 'non-overlapping' και την 'overlapping'. Στην πρώτη περίπτωση, οι χρονοσειρές χωρίζονται σε διαδοχικές μη επικαλυπτόμενες χρονικές περιοχές όπου το μήκος της κάθε περιοχής ισούται με το επίπεδο συνάθροισης. Αντίθετα, η δεύτερη περίπτωση είναι παρόμοια με την τεχνική του κινούμενου παραθύρου (moving window) όπου το μέγεθος του παραθύρου είναι ίσο με το επίπεδο συνάθροισης. Σε κάθε περίοδο, το παράθυρο κινείται ένα βήμα μπροστά, έτσι ώστε η παλαιότερη παρατήρηση να βγει εκτός και να συμπεριληφθεί η νεότερη.

3.2.1. Μη Επικαλυπτόμενη Χρονική Συνάθροιση

Χρησιμοποιώντας την τεχνική της χρονικής συνάθροισης (temporal aggregation) μετατρέπεται μια υψηλής συχνότητας χρονοσειρά σε μια χαμηλής συχνότητας. Με την τεχνική αυτή επιτυγχάνεται καλύτερη εκτίμηση των χαρακτηριστικών κάθε χρονοσειράς καθώς και μείωση της επίδρασης του μοντέλου πρόβλεψης που χρησιμοποιείται. Σε χαμηλά επίπεδα συνάθροισης φαίνονται περισσότερο περιοδικά χαρακτηριστικά της χρονοσειράς, όπως η εποχικότητα. Όσο αυξάνονται τα επίπεδα συνάθροισης δίνεται περισσότερη βάση σε χαμηλόσυχνα χαρακτηριστικά της χρονοσειράς, όπως στο επίπεδο και στη τάση της χρονοσειράς.

Στόχος είναι η ταυτοποίηση, η επιλογή και ο υπολογισμός των παραμέτρων των μοντέλων πρόβλεψης.



Εικόνα 11. Αποτύπωση μη-επικαλυπτόμενης χρονικής συνάθροισης.

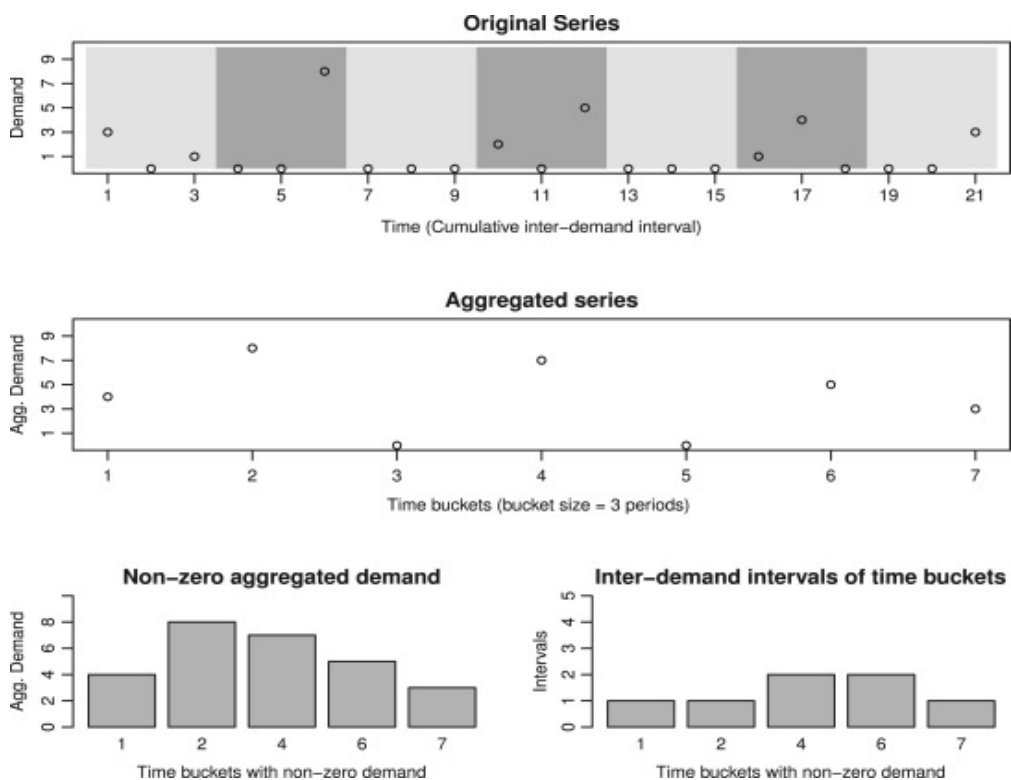
Αφού γίνει η χρονική συνάθροιση, παράγονται προβλέψεις για κάθε επίπεδο συνάθροισης χρησιμοποιώντας ένα κατάλληλο μοντέλο πρόβλεψης. Ο B. Abraham (1982) χρησιμοποίησε την τεχνική της χρονικής συνάθροισης (temporal aggregation) για να πραγματοποιήσει προβλέψεις, θέλοντας να δει την επίδραση της μεθόδου στην αποτελεσματικότητα των προβλέψεων αλλά και την επίδραση των αρχικών δεδομένων που χρησιμοποιούμε. Τα συμπεράσματα του δεν μπορεί να είναι απόλυτα καθώς χρησιμοποιήθηκε μικρό πλήθος χρονοσειρών, ωστόσο κατέληξε στο συμπέρασμα ότι η χρονική συνάθροιση επιδρά θετικά στην αποτελεσματικότητα του μοντέλου πρόβλεψης (ARIMA).

Οι M.Z. Babai και M.M. Ali (2012) εφάρμοσαν την μέθοδο της χρονικής συνάθροισης σε δεδομένα με διακοπτόμενη ζήτηση. Η διαχείριση της διακοπτόμενης ζήτησης και η πραγματοποίηση των προβλέψεων είναι αρκετά δύσκολη λόγω της παρουσίας διαστημάτων μηδενικών τιμών. Ένας τρόπος διαχείρισης της δυσκολίας αυτής και της πραγματοποίησης των προβλέψεων είναι η χρήση της μεθόδου συνάθροισης. Τα οφέλη της μεθόδου προέκυψαν εμπειρικά από τις προβλέψεις που πραγματοποιήθηκαν για δεδομένα της Royal Air Force (UK). Τα αποτελέσματα της μελέτης δείχνουν ότι η χρήση της συνάθροισης οδηγεί σε μεγαλύτερη ακρίβεια από την κλασική μέθοδο πρόβλεψης. Με την χρήση της συνάθροισης μειώθηκαν ή και εξαφανίστηκαν εντελώς οι περίοδοι μηδενικών τιμών, ενώ περισσότερα μοντέλα

προβλέψεων έγιναν διαθέσιμα για χρήση και τέλος, δεν υπάρχει λόγος να ξαναγίνει χωρισμός των προβλέψεων γιατί στην συγκεκριμένη περίπτωση ενδιαφερόμαστε για την συνολική πρόβλεψη.

3.2.2. Επικαλυπτόμενη Χρονική Συνάθροιση (

Όσον αφορά στην τεχνική της επικαλυπτόμενης χρονικής συνάθροισης (overlapping temporal aggregation) έχουν πραγματοποιηθεί λίγες μελέτες για την επίδραση της. Οι Κ.Η. Luiz και Α.Μ. Pedro (1992) ανέλυσαν τις επιδράσεις της σε μοντέλα ARIMA και κατέληξαν στο συμπέρασμα ότι το μοντέλο από ARIMA (p,d,q) μετατρέπεται σε ARIMA (P,d,Q). Αντίστοιχες μελέτες πραγματοποιήθηκαν και από τους Κ. Vrannas και J. Hellstrom (2002) σε μοντέλα INARMA (1, 0, 0) που παρέμειναν τα ίδια, καθώς και από τους Μ. Mohammadpour και J.E. Boylan (2012) σε μοντέλα INARMA (p, 0, q) όπου και αυτά παρέμειναν τα ίδια.



Εικόνα 12. Απεικόνιση μη-επικαλυπτόμενης χρονικής συνάθροισης.

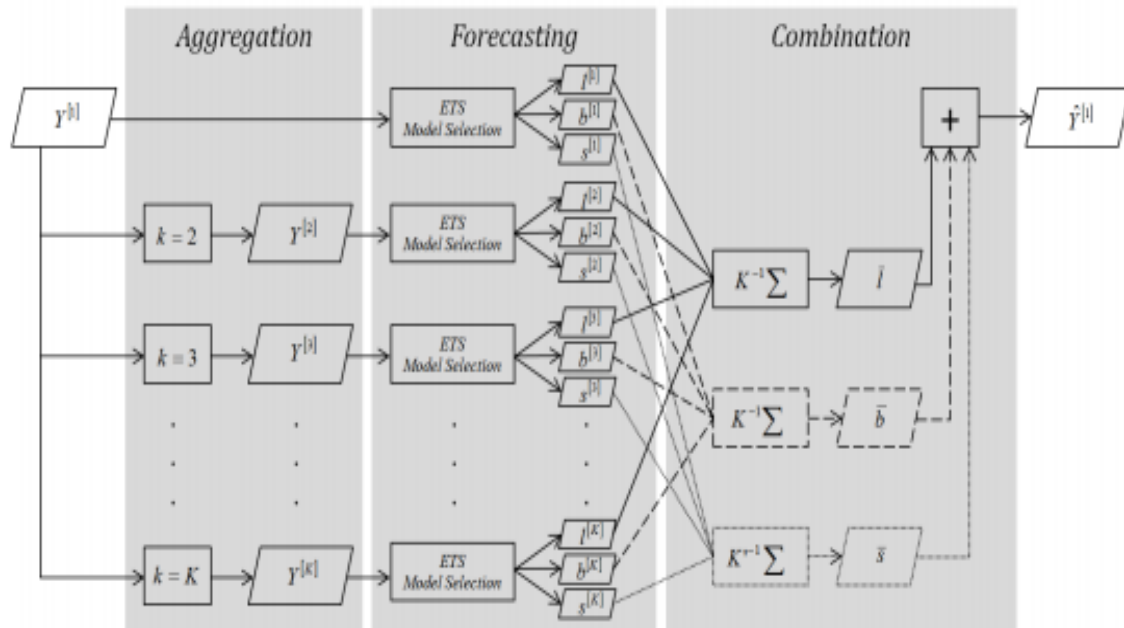
3.2.3. Αλγόριθμος προβλέψεων πολλαπλής συνάθροισης

Η μεθοδολογία MAPA (Multiple Aggregation Prediction Algorithm) επεκτείνει την μέθοδο της μη επικαλυπτόμενης συνάθροισης και τον συνδυασμό των προβλέψεων. Έχει προταθεί από τους N. Kourentzes, F. Petropoulos και J. Trapero (2014). Η φιλοσοφία της μεθόδου είναι η χρησιμοποίηση πολλών επιπέδων συνάθροισης για την υπό-μελέτη χρονοσειρά, τονίζοντας κάθε φορά διαφορετικό ποιοτικό χαρακτηριστικό της. Στη συνέχεια πραγματοποιούνται οι προβλέψεις για το κάθε επίπεδο συνάθροισης χρησιμοποιώντας κατάλληλη μέθοδο πρόβλεψης. Σε χαμηλά επίπεδα συνάθροισης τονίζονται χαρακτηριστικά όπως η εποχικότητα ενώ όσο αυξάνεται το επίπεδο τονίζονται χαρακτηριστικά όπως είναι η τάση και το επίπεδο της χρονοσειράς. Στην συνέχεια, αντί να συνδυάζονται απευθείας οι προβλέψεις από τα διαφορετικά επίπεδα συνάθροισης, χωρίζονται οι χρονοσειρές ανάλογα με τα ποιοτικά χαρακτηριστικά τους και συνδυάζονται. Στο σημείο αυτό πρέπει να τονιστεί πως ο συνδυασμός που θα γίνει θα πρέπει να περιέχει τα ποιοτικά χαρακτηριστικά που έχουν επιλεγεί, ενώ το επιλεγμένο επίπεδο συνάθροισης δεν πρέπει να οδηγεί σε απώλεια δεδομένων. Αναφέρεται χαρακτηριστικά, πως εάν συνδυαστούν μία εποχικοποιημένη πρόβλεψη με μια μη εποχικοποιημένη, το αποτέλεσμα θα είναι μία πρόβλεψη που θα έχει την μισή εποχικότητα από την αρχική.

Με την μέθοδο MAPA επιτυγχάνονται τα εξής:

- μείωση της επίδρασης της μεθόδου που χρησιμοποιείται για την πρόβλεψη,
- συνδυασμός των πλεονεκτημάτων ακριβείας από την συνάθροιση και από την συνδυασμένη πρόβλεψη
- μεγαλύτερη ακρίβεια.

Η εφαρμογή της μεθόδου είχε αρκετά πλεονεκτήματα στις μεθόδους Naive, SES και Theta, ωστόσο δεν έχει μελετηθεί πλήρως η επίπτωση της στα διάφορα μοντέλα πρόβλεψης χρονοσειρών. Θεωρείται μία επιτυχημένη μέθοδος πρόβλεψης καθώς βελτιώνει την ακρίβεια σε περιπτώσεις χρονοσειρών μεγάλης τυχαιότητας, ειδικότερα σε περιπτώσεις μακροπρόθεσμης πρόβλεψης. Κρίσιμος παράγοντας στην επιτυχία της μεθόδου είναι η χρήση πολλών και όχι ενός συγκεκριμένου επιπέδου της συνάθροισης για την παραγωγή προβλέψεων. Αυτό είναι αρκετά βοηθητικό δεδομένου ότι δεν υπάρχει κάποιο οικουμενικά βέλτιστο επίπεδο συνάθροισης για κάθε χρονοσειρά.



Εικόνα 13. Απεικόνιση της μεθοδολογίας MAPA.

4. Πειραματική Διαδικασία

4.1. Στόχος της Πειραματικής Διαδικασίας

Σε προηγούμενο κεφάλαιο πλαισιώθηκε η έννοια των κρυπτονομισμάτων, η λειτουργία και η χρησιμότητάς τους, αλλά και η προσέγγισή τους ως επενδυτική επιλογή τόσο για χρηματοπιστωτικά ιδρύματα όσο και από μικρούς επενδυτές. Παράλληλα, παρουσιάστηκαν κάποιες βασικές αρχές και μέθοδοι σύμφωνα με τις οποίες οι επενδυτές λαμβάνουν αποφάσεις για να καθορίσουν τις επόμενες κινήσεις τους στην αγορά. Ο στόχος της παρούσας εργασίας, όπως προαναφέρθηκε, είναι η δημιουργία ενός μοντέλου που θα χρησιμοποιείται ως επενδυτικό εργαλείο που συμβάλλει στο αποτελεσματικότερο trading σε binary option βασισμένα σε κρυπτονομίσματα. Παρακάτω μελετάται ένας συνδυασμός μεθόδων για την υλοποίηση του μοντέλου, τόσο ορισμένων από τις κλασικές μεθόδους πρόβλεψης που αναφέρθηκαν στο Κεφάλαιο 2, όσο και της Χρονικής Συνάθροισης που παρουσιάστηκε στο Κεφάλαιο 4.

4.2. Επιλογή Επενδυτικού Προϊόντος

Στο πλαίσιο της διπλωματικής εργασίας ως επενδυτικό αγαθό χρησιμοποιήθηκε το Συμβόλαιο επί Διαφοράς (CFD) του κρυπτονομίσματος bitcoin. Τα CFDs είναι παράγωγα διαπραγματεύσιμα προϊόντα που παρέχουν ευκαιρίες για συναλλαγές στην κίνηση των τιμών από πλήθος χρηματοοικονομικών προϊόντων όπως δείκτες μετοχών και συμβόλαια εμπορευμάτων.

Ειδικότερα, στην περίπτωση των κρυπτονομισμάτων παρέχεται η δυνατότητα άμεσων συναλλαγών με το παράγωγο αυτό κλειδωμένο στην τιμή του κρυπτονομίσματος που επιλέγει ο επενδυτής χωρίς ουσιαστικά να κατέχει το νόμισμα. Αυτό έπεται ότι ουσιαστικά δεν μπορεί να εξαργηθεί η ανταλλακτική αξία του κρυπτονομίσματος για την αγορά ενός αγαθού ή υπηρεσίας, παρά μόνο να αγορασθεί ή να πωληθεί σαν επενδυτικό προϊόν. Στο πλαίσιο αυτό όμως δεν υπάρχει η χρονική καθυστέρηση της επιβεβαίωσης του blockchain, ούτε τα τέλη συναλλαγών (trading fees) που περιλαμβάνει η αγοραπωλησία πραγματικών κρυπτονομισμάτων. Αξίζει να υπογραμμισθεί ότι τα μόνα έξοδα συναλλαγής είναι το spread, που κυμένεται από 0.6% έως 0.7%, στο eToro (<https://www.etoro.com>), ένα από τα μεγαλύτερα συναλλακτήρια CFD της Ευρώπης. Βάση αυτών των κριτηρίων επιλέχθηκε το CFD του κρυπτονομίσματος Bitcoin ως το καταλληλότερο επενδυτικό αγαθό για την εκπόνηση της παρούσας διπλωματικής εργασίας.

4.3. Συλλογή και Εξόρυξη Δεδομένων

Για τη δημιουργία του κατάλληλου μοντέλου πρόβλεψης της τιμής του παραγώγου CFD του κρυπτονομίσματος Bitcoin ήταν αναγκαία η περισυλλογή ιστορικών δεδομένων των τιμών αυτών. Η χρησιμότητα αυτού έγκειται στο να

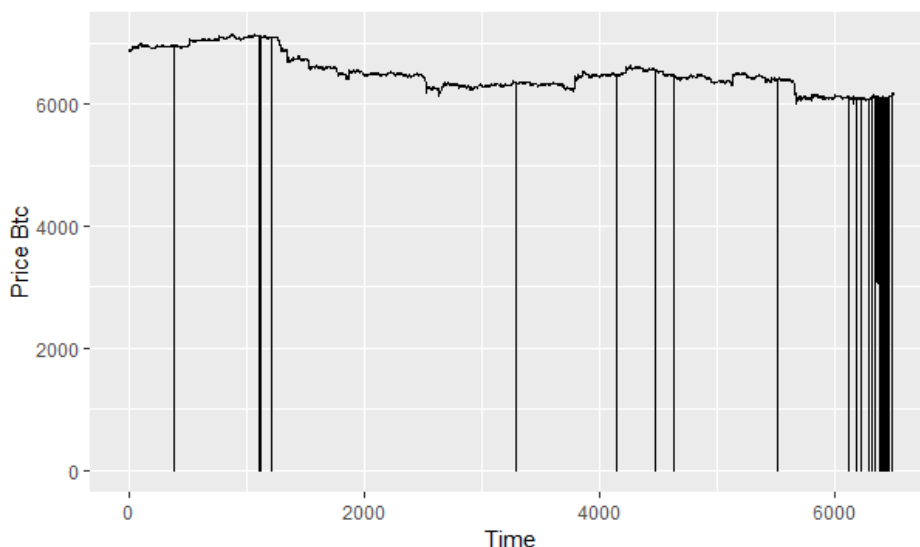
αποφανθεί ότι η χρήση της μεθόδου που επιλέχθηκε θα είχε τα καλύτερα αποτελέσματα εάν είχε χρησιμοποιηθεί από κάποιον επενδυτή στο παρελθόν. Καθώς η τιμή του CFD είναι κλειδωμένη στην πραγματική τιμή του κρυπτονομίσματος, η ανάλυση και μελέτη των ιστορικών τιμών του bitcoin αποτελούν το βασικό άξονα για την παραγωγή προβλέψεων. Πρόκειται για δεδομένα σε επίπεδο λεπτού του ζεύγους BTC/USD (bitcoin/United State Dollar) που εκφράζουν την τιμή αγοράς ενός bitcoin, συναρτήσει του χρόνου. Τα δεδομένα αυτά αφορούν σχεδόν 5 μήνες, από την 1/8/2018 έως τις 17/12/2018. Τα ιστορικά δεδομένα συλλέχθηκαν από τις ιστοσελίδες «bitcoinaverage.com» και «cryptoarchive», όπου ζητήθηκαν και παραχωρήθηκαν για ερευνητικό σκοπό ως raw data ανά ημέρα σε αρχεία με μορφή “.csv”.

4.4. Διαχείριση Ασυνήθιστων Τιμών

Τα δεδομένα, μέσω προγράμματος Microsoft Excel, επεξεργάστηκαν για πρώτη φορά και συγχωνεύτηκαν σε 5 αρχεία τύπου csv που το κάθε ένα περιλάμβανε τις παρατηρήσεις για κάθε μήνα. Παρατηρήθηκε η εμφάνιση ορισμένων συμβόλων (τύπου ^, \$, /, *) κατ' επανάληψη σε μεγάλο αριθμό κελιών για άγνωστο λόγο, πιθανώς κακόβουλα. Μέσω της διαδικασίας αναζήτησης και αντικατάστασης του Excel, επεξεργάστηκαν τα αρχεία με αποτέλεσμα να παραμείνουν μόνο αριθμητικοί χαρακτήρες με δύο το πολύ δεκαδικά ψηφία, τα δεδομένα που περιέχουν ουσιαστικά τα προς ανάλυση ποιοτικά χαρακτηριστικά.

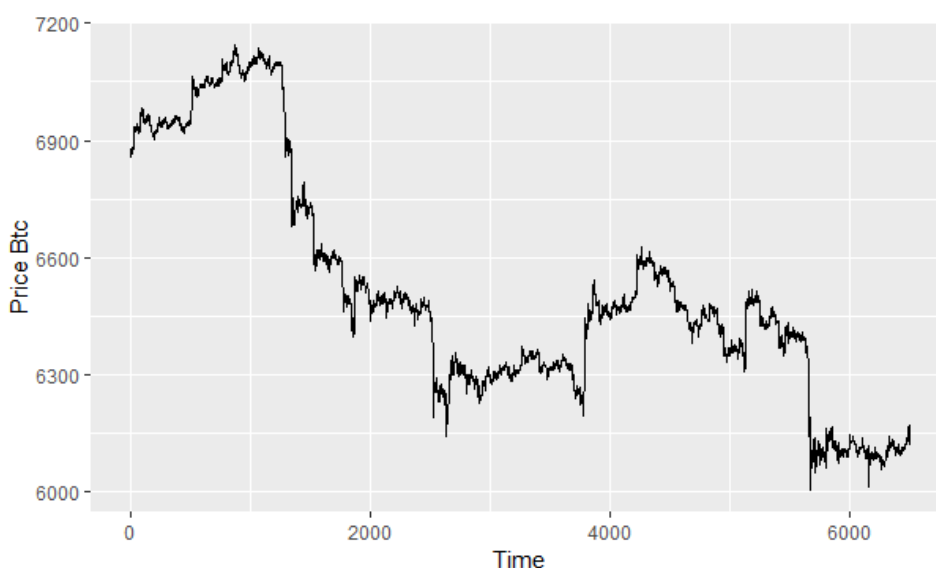
Στη συνέχεια τα δεδομένα εισήχθησαν στο περιβάλλον προγραμματισμού RStudio της γλώσσας προγραμματισμού R, όπου και εκπονήθηκε το μεγαλύτερο μέρος της παρούσας διπλωματικής. Η διαδικασία αυτή υλοποιήθηκε με την συνάρτηση “read_excel” και για τη συγχώνευση των παρατηρήσεων σε ένα data frame χρησιμοποιήθηκε η εντολή “rbind()”. Στο Παράρτημα της παρούσας υπάρχουν τόσο οι δύο προαναφερθείσες συναρτήσεις όσο και η υλοποίηση της χρονικής ακολουθίας (timestamp) που χρησιμοποιήθηκε για να μετατρέψουμε τις παρατηρήσεις μας σε χρονοσειρά.

Αξίζει να σημειωθεί ότι καθώς τα δεδομένα βρισκόταν σε επίπεδο λεπτού και είχαν εξορυχθεί από συναλλακτήρια αξιοποιώντας την τεχνολογία REST API, περιείχαν αρκετές μηδενικές τιμές, είτε λόγο προσωρινής διακοπής της σύνδεσης στο διαδίκτυο είτε λόγω τεχνικού προβλήματος του συναλλακτηρίου. Επίσης, πραγματοποιήθηκε έλεγχος για την εύρεση εσφαλμένων τιμών, αρνητικών τιμών ή με εξαιρετική απόκλιση από τον υπόλοιπο πληθυσμό (<99,75%), που θα μπορούσαν να προκαλέσουν προβλήματα και αστοχίες τόσο στην εκτέλεση του κώδικα, όσο και στην ποιότητα των αποτελεσμάτων. Έτσι, κρίθηκε αναγκαία η αντιμετώπιση όλων των παραπάνω ασυνήθιστων τιμών.



Εικόνα 14. Χαρακτηριστικό παράδειγμα ασυνήθιστων τιμών πριν την αντικατάστασή τους.

Η διαδικασία που ακολουθήθηκε για την αντικατάσταση των παραπάνω τιμών ξεκίνησε με την μετατροπή όλων των αυτών που ήταν μικρότερες από 2000 \$ και μεγαλύτερες από 9000\$ σε Na τιμές. Έπειτα, όλες οι Na τιμές αντικαταστάθηκαν με τις πιο πιθανές τιμές που θα μπορούσαν να λάβουν μέσω της συνάρτησης “na.kalman”, με default arguments. Περισσότερες πληροφορίες για τη διαδικασία που χρησιμοποιεί η kalman παρουσιάζονται στο Παράρτημα. Λόγω της σποραδικότητας αυτών των περιπτώσεων, κρίνεται ότι η τόσο η διαδικασία επιλογής του μοντέλου των προβλέψεων, όσο και ο υπολογισμός του τελικού σφάλματος θα μείνουν πρακτικά ανεπηρέαστα.



Εικόνα 15. Χαρακτηριστικό παράδειγμα ασυνήθιστων τιμών μετά την αντικατάστασή τους.

4.5. Επιλογή Μεθόδου Πρόβλεψης

Το επόμενο στάδιο για την παραγωγή των καλύτερων προβλέψεων της τιμής κλεισίματος του Btc/Usd ήταν η αξιολόγηση των μεθόδων πρόβλεψης και η επιλογή εκείνου του μοντέλου με το μικρότερο σφάλμα. Ως σφάλμα αξιολόγησης επιλέχθηκε το Μέσο Απόλυτο Σφάλμα (Mean Absolute Error). Πιο αναλυτικά, οι μέθοδοι που εξετάστηκαν ήταν οι εξής :

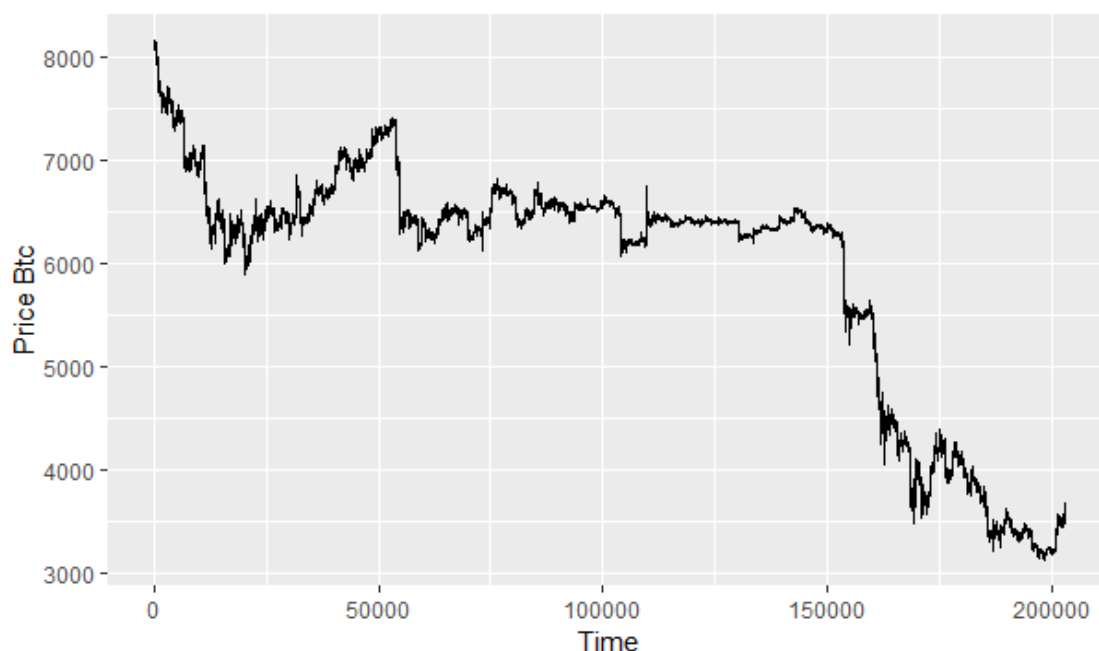
- Naïve
- Holt
- Simple Exponential Smoothing (SES)
- Arima
- Απλός Κινητός Μέσος Όρος

Ο συνολικός πληθυσμός τιμών είναι 203,040, ενώ το δείγμα που επιλέχθηκε για την αξιολόγηση των προβλέψεων είναι από την τιμή 100,000 έως την τιμή 150,000. Ή σε πραγματικό χρόνο από 9/10/2018 10:30 έως 13/11/2018 14:20. Σε αυτό το δείγμα πραγματοποιήθηκαν οι δοκιμές για κάθε μία μέθοδο πρόβλεψης ξεχωριστά, με ορίζοντα 60, και με παρατηρήσεις υλοποίησης 15, 60 και 120.

Ο ορίζοντας επιλέχθηκε με αυτόν τον τρόπο, καθώς ο στόχος είναι η ωριαία πρόβλεψη στο κλείσιμο, δηλαδή η τελευταία από τις 60 παρατηρήσεις της πρόβλεψης. Με βάση το μέσο απόλυτο σφάλμα της κάθε μεθόδου, επιλέχθηκε η "SES" ως η καταλληλότερη για τα ποιοτικά χαρακτηριστικά της παρούσας χρονοσειράς. Τα αποτελέσματα παρατίθενται αναλυτικά στο Κεφάλαιο 5, «Παρουσίαση Αποτελεσμάτων» και συγκεκριμένα στο Υποκεφάλαιο 5.1.

4.6. Χρονική Συνάθροιση της χρονοσειράς (Temporal Aggregation)

Στη συνέχεια , όπως παρατηρήθηκε οπτικά από την απεικόνιση της χρονοσειράς, είναι εμφανές ότι αυτή περιέχει σε σχετικά μεγάλο βαθμό τυχαιότητα (randomness). Επίσης όπως διαφαίνεται παρακάτω υπάρχει και μία αρνητική τάση (trend) στη χρονοσειρά, καθώς τα δεδομένα συλλέχθηκαν λίγους μήνες αφού η Κίνα είχε αποκλείσει νομικά τη διαδικασία της εξόρυξης (mining) bitcoin , σε μία προσπάθεια να σταματήσει την εκθετική ανάπτυξη των κρυπτονομισμάτων στη χώρα, και υπήρχε μία γενική ανησυχία στην αγορά σχετικά με το μέλλον του κλάδου. Παρακάτω παρατίθεται μία απεικόνιση όλων των ιστορικών τιμών που συλλέχθηκαν για να γίνουν καλύτερα αντιληπτά όσα ειπώθηκαν προηγουμένως.

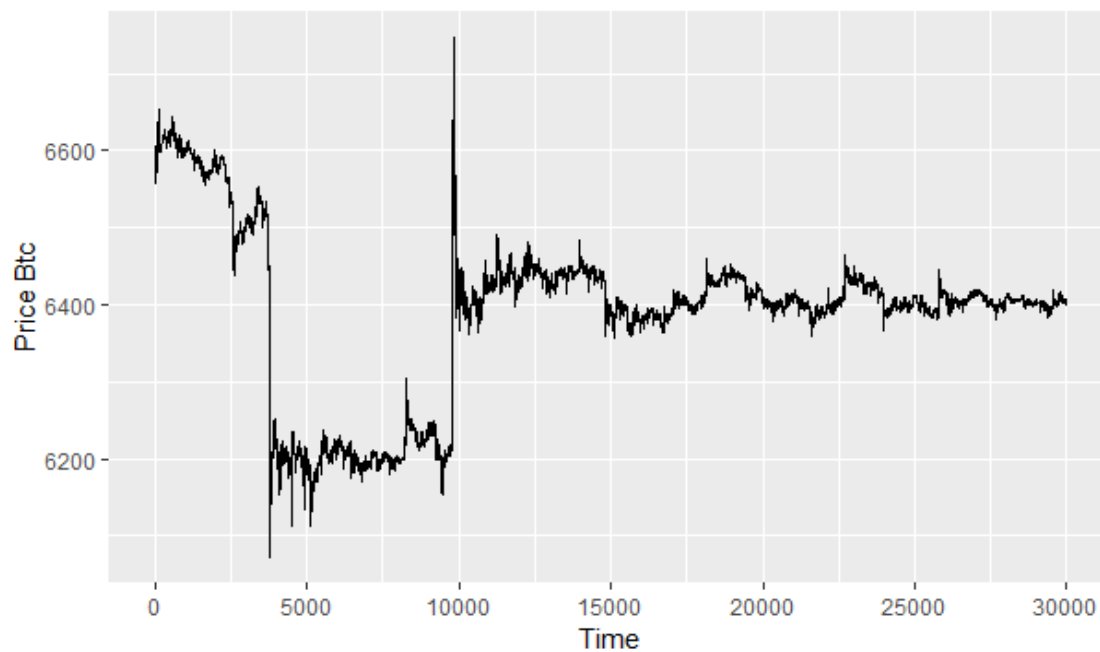


Εικόνα 16. Απεικόνιση όλων των δεδομένων που συλλέχθηκαν.

Στο πλαίσιο αυτής της διπλωματικής για την εξάλειψη της τυχαιότητας και την προσπάθεια επίτευξης ακριβέστερων αποτελεσμάτων εξετάστηκε η αξιοποίηση της τεχνικής της χρονικής συνάθροισης. Ουσιαστικά δημιουργήθηκαν 5 επίπεδα ιεράρχησης από υψηλή συχνότητα σε επίπεδο λεπτού, μέχρι πιο χαμηλή συχνότητα σε επίπεδο ώρας, με μη επικαλυπτόμενα (non-overlapping) επίπεδα. Πιο αναλυτικά, δημιουργήθηκαν τα set1, set5, set15, set30 και set60 τα οποία περιλαμβάνουν τα δεδομένα σε επίπεδο λεπτού και συναθροισμένα (aggregated) σε επίπεδο 5λεπτού, 15λεπτού, μισής ώρας και ωριαία αντίστοιχα. Το κάθε επίπεδο χρονικής συνάθροισης δημιουργήθηκε από την αρχική χρονοσειρά, παίρνοντας το μέσο όρο τόσων τιμών όσων και των λεπτών της συχνότητας που περιλαμβάνει. Για παράδειγμα,

$$Set30[1] = \frac{1}{30} * \sum_{i=1}^{30} Set1[i].$$

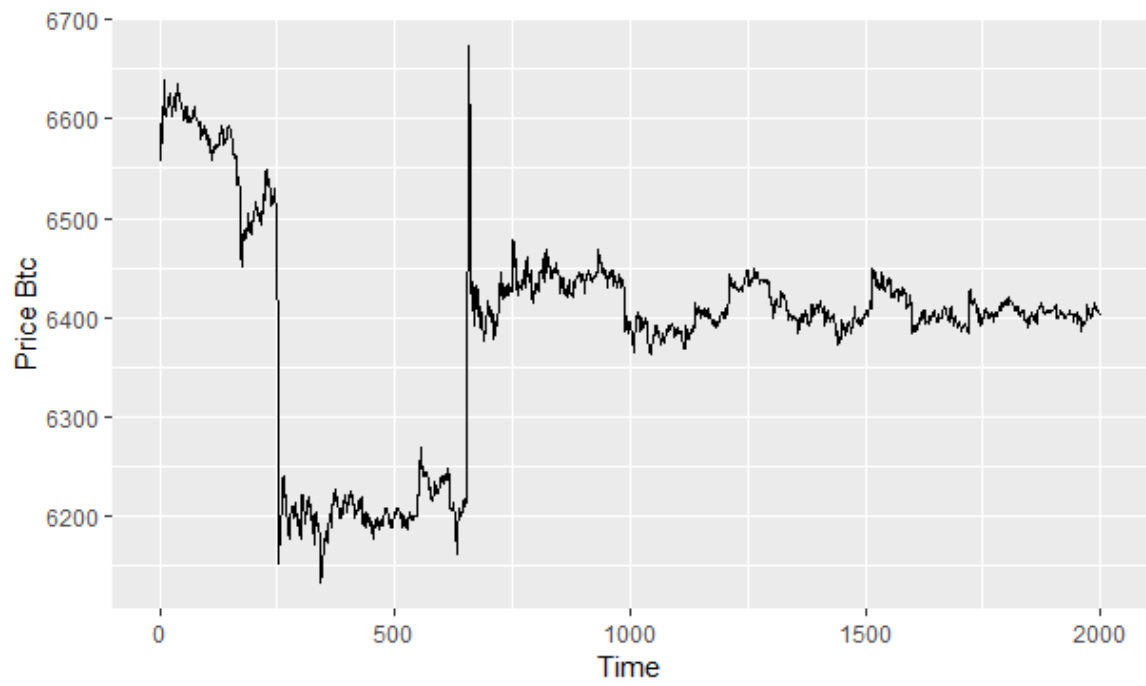
Παρακάτω παρατίθενται στιγμιότυπα αποτύπωσης των προαναφερθέντων επιπέδων συνάθροισης. Τα στιγμιότυπα αφορούν στον ίδιο πραγματικό χρόνο, ώστε να διευκολύνεται οπτικά η διαφοροποίησή τους. Το χρονικό διάστημα που αποτυπώνεται αποτελεί μέρος του δείγματος που χρησιμοποιήθηκε για τις προβλέψεις.



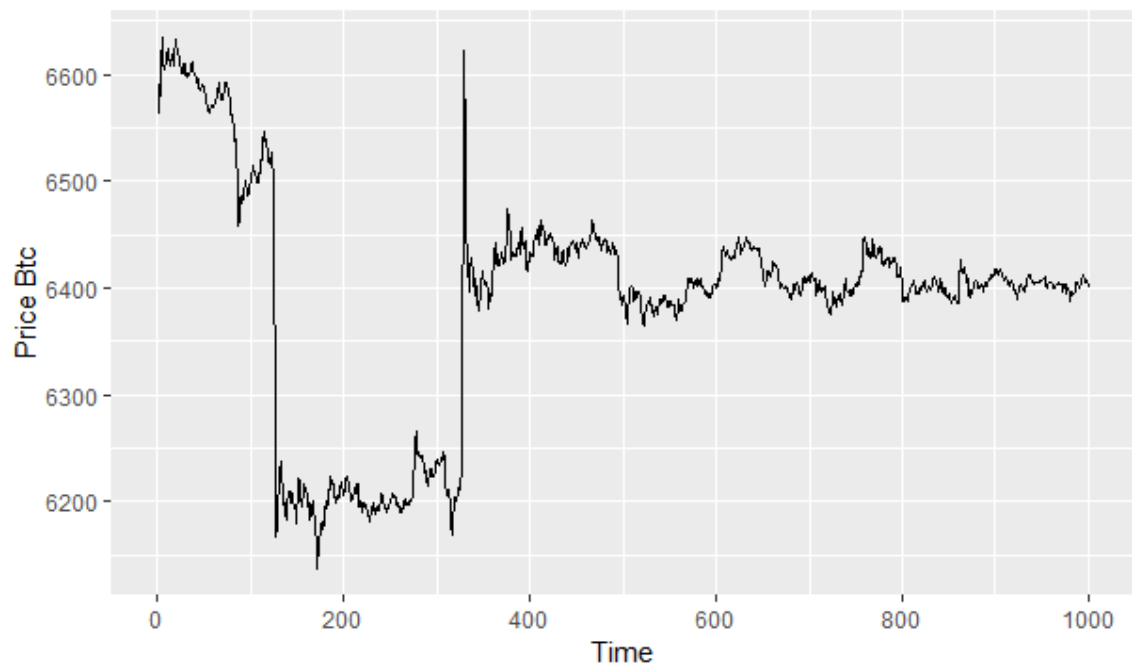
Εικόνα 17. Set 1 min



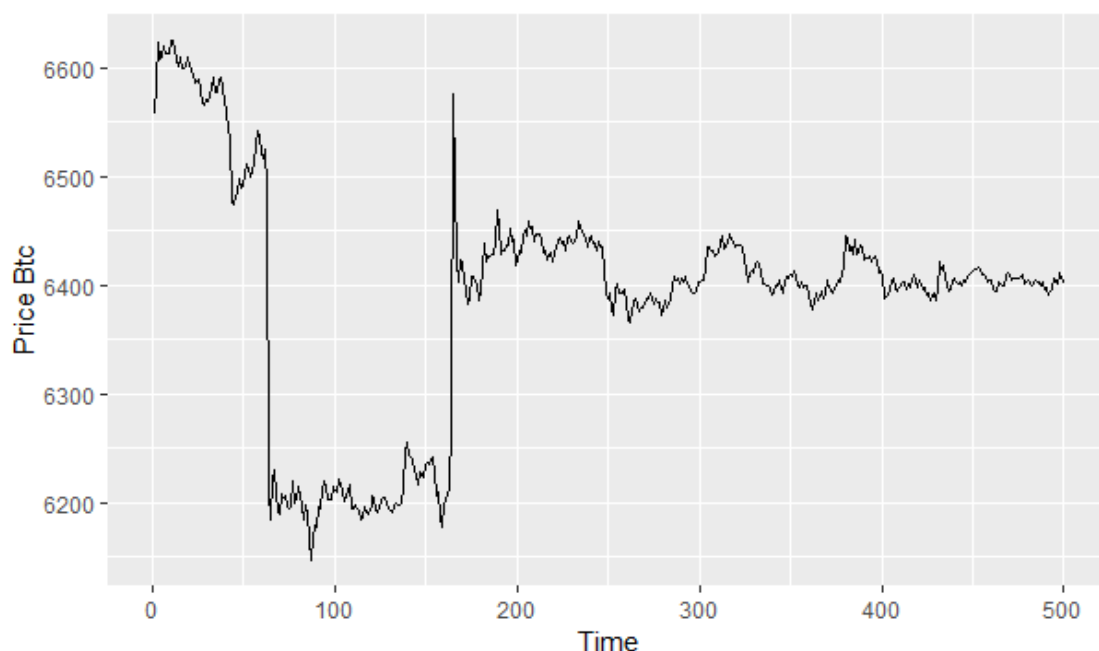
Εικόνα 18. Set 5 mins



Εικόνα 19. Set 15 mins



Εικόνα 20. Set 30 mins



Εικόνα 21. Set 60 mins

Είναι εμφανές ότι περισσότερο αυξάνουμε το επίπεδο χρονικής συνάθροισης , τόσο η χρονοσειρά μας γίνεται πιο καθαρή , με διακριτή την εξομάλυνση της τυχαιότητας. Γεγονός που καθιστά τις προβλέψεις πιο εύκολες από αλγοριθμική σκοπιά αλλά πιθανώς και περισσότερο ακριβείς.

4.7. Παραγωγή προβλέψεων σε διαφορετικά επίπεδα χρονικής συνάθροισης

Τα διαφορετικά επίπεδα χρονικής συνάθροισης δημιουργήθηκαν με σκοπό να παραχθούν προβλέψεις στο κάθε ένα ξεχωριστά και να εξετασθεί η ακρίβειά τους. Αξίζει να υπογραμμισθεί ότι το δείγμα το οποίο χρησιμοποιήθηκε για την αξιολόγηση των μεθόδων πρόβλεψης είναι το ίδιο πάνω στο οποίο πραγματοποιήθηκαν προβλέψεις για τα διάφορα επίπεδα χρονικής συνάθροισης. Αυτό εκπηγάει από την ανάγκη να έχουμε ένα δείγμα αναφοράς και τα αποτελέσματα μεταξύ των διαφορετικών μεθόδων και τεχνικών προβλέψεων να είναι συγκρίσιμα. Σε ότι αφορά τη χρονική συνάθροιση, στο εκάστοτε επίπεδο ιεράρχησης χρησιμοποιήθηκαν παράμετροι στις προβλέψεις προσαρμοσμένες ανάλογα με τη συχνότητα της εκάστοτε χρονοσειράς.

Πιο συγκεκριμένα, για την ωριαία πρόβλεψη χρησιμοποιήθηκαν ορίζοντες 12 , 4 , 2 και 1 για τα set5, set15, set30 και set60 αντίστοιχα. Σε κάθε επίπεδο πραγματοποιήθηκαν προβλέψεις με 4 διαφορετικές παρατηρήσεις υλοποίησης 12, 24, 48 , 96 και 192 (αναλυτικότερα στο Κεφάλαιο 5) και υπολογίσθηκαν τα αντίστοιχα σφάλματα. Οι παρατηρήσεις υλοποίησης είναι λιγότερες από αυτές που χρησιμοποιήθηκαν στην αρχική αξιολόγηση των μεθόδων πρόβλεψης καθώς για τα επίπεδα με χαμηλότερες συχνότητες δεν θα είχε νόημα να εξετάσουμε τόσα πολλά

δεδομένα. Όλες οι προβλέψεις καθώς και τα απόλυτα μέσα σφάλματα αυτών αποθηκεύτηκαν σε ξεχωριστούς πίνακες και παρουσιάζονται αναλυτικά στο Κεφάλαιο 5.

4.8. Συνδυασμός διαφορετικών επιπέδων χρονικής συνάθροισης

Στη συνέχεια, στην προσπάθεια υλοποίησης αποδοτικότερων προβλέψεων εξετάστηκε η εφαρμογή του Multiple Aggregation Prediction Algorithm (MAPA), δηλαδή του συνδυασμού των προβλέψεων κάθε διαφορετικού επιπέδου χρονικής ιεράρχησης για την παραγωγή της τελικής πρόβλεψης.

Αρχικά, η πρώτη ιδέα ήταν να χρησιμοποιηθεί ο μέσος όρος των προβλέψεων διαφορετικών συχνοτήτων για την πρόβλεψη της κάθε επόμενης ώρας. Με γνώμονα αυτό εκτελέστηκε από την αρχή ο αλγόριθμος με τα κατάλληλα βάρη σε κάθε βαθμίδα και παρήχθησαν οι προβλέψεις για ολόκληρο το δείγμα. Το σφάλμα που προέκυψε είναι ελαφρώς βελτιωμένο αλλά είναι εμφανές ότι γεννάται το ερώτημα, εάν το σφάλμα θα μειωθεί περαιτέρω με καταλληλότερα βάρη.

Έτσι για την εύρεση των κατάλληλων βαρών και καθώς η συνάρτηση είναι γραμμική, χρησιμοποιήθηκε η συνάρτηση *optim* της γλώσσας προγραμματισμού R. Η συνάρτηση *optim* χρησιμοποιείται με σκοπό την εύρεση των παραμέτρων μίας γραμμικής συνάρτησης, μεγιστοποιώντας ή ελαχιστοποιώντας την τιμή της. Στη δεδομένη περίπτωση χρησιμοποιήθηκε για την ελαχιστοποίηση της αντικειμενικής συνάρτησης:

$$F_{optim} = \text{mean}(\text{abs}(\text{FinForecast} - \text{Price}))$$

$$\text{Όπου, } \text{FinForecast} = w_1 * \text{Forecast}_{set1} + w_2 * \text{Forecast}_{set5} + w_3 * \text{Forecast}_{set15} + w_4 * \text{Forecast}_{set30} + w_5 * \text{Forecast}_{set60}$$

Με $\text{Forecast}_{set,i}$, $i = 1, 5, 15, 30, 60$ να είναι το άθροισμα όλων των προβλέψεων που παρήχθησαν για το δείγμα από το αντίστοιχο set.

Όπου, $w_i, i=1, 2, 3, 4, 5$ με $0 \leq w_i \leq 1$ και

$$\sum_{i=1}^5 w_i = 1$$

είναι τα βάρη για κάθε επίπεδο χρονικής συνάθροισης.

Τα αποτελέσματα των βαρών μετά την εκτέλεση της συνάρτησης *optim* είναι:

$$w_1 = 0.4425617 \quad w_2 = 0.00002386825 \quad w_3 = 0.2746586$$

$$w_4 = 0.199256 \quad w_5 = 0.08349906$$

Όπως είναι εμφανές, είναι αναγκαίο να επαληθευτεί σε όλο το δείγμα ότι το άθροισμα των προβλέψεων των 5 διαφορετικών επιπέδων χρονικής συνάθροισης

πολλαπλασιασμένο με το εκάστοτε βάρος του δίνει καλύτερα αποτελέσματα σε ολόκληρο το δείγμα και όχι μόνο στο άθροισμα όλων των προβλέψεων. Συνεπώς υπολογίσθηκαν εξ αρχής προβλέψεις για όλο το εύρος του δείγματος με τον τρόπο που περιγράφεται παραπάνω και βρέθηκε το νέο απόλυτο μέσο σφάλμα.

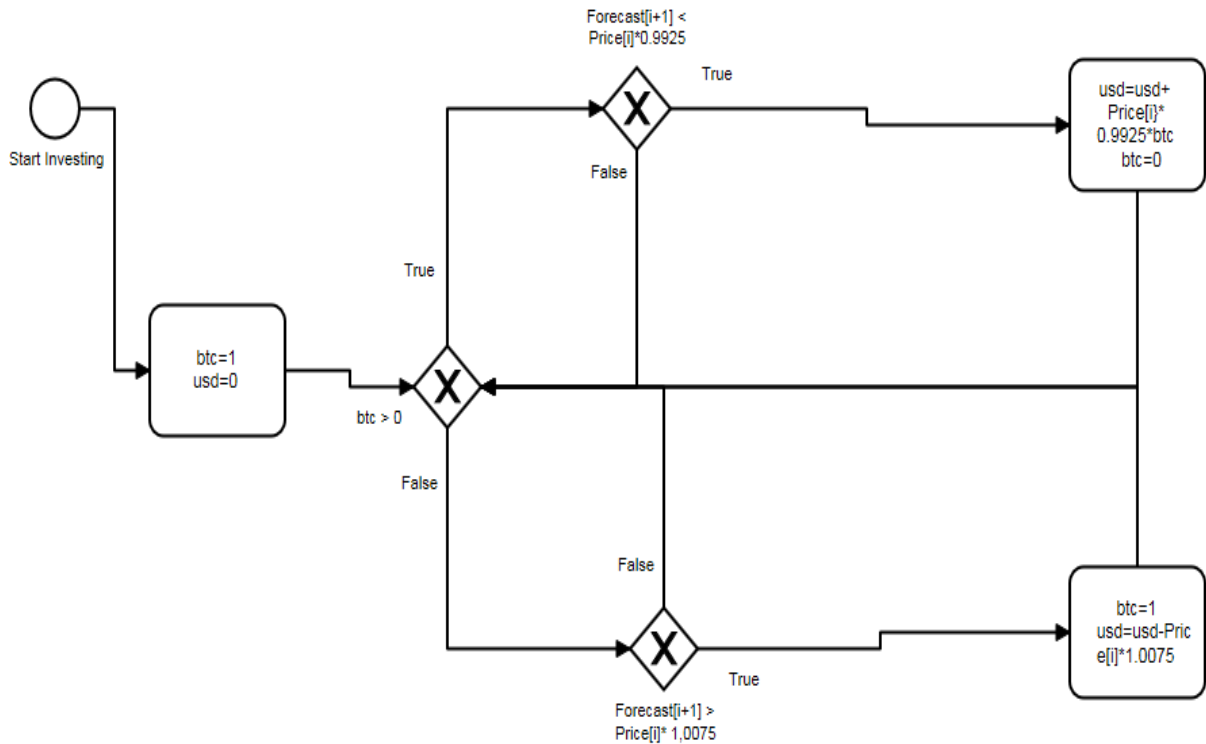
4.9. Χρήση του μοντέλου πρόβλεψης ως εργαλείου αγοροπωλησίας

Λαμβάνοντας υπόψιν όλα τα παραπάνω, το μοντέλο που δημιουργήθηκε περιλαμβάνει και τα 5 επίπεδα χρονικής συνάθροισης, με τα οποία δημιουργούμε 5 προβλέψεις με ορίζοντα 60, 12, 4, 2, 1 τιμές για τα set 1, 5, 15, 30 και 60 αντίστοιχα. Οι 5 αυτές προβλέψεις πολλαπλασιάζονται με τα βάρη: w_1, w_2, w_3, w_4 και w_5 . Το άθροισμα αυτών αποτελεί την ακριβέστερη πρόβλεψη σύμφωνα με την ανάλυση της παρούσας διπλωματικής εργασίας.

Το παραπάνω μοντέλο χρησιμοποιήθηκε ως εργαλείο για τη λήψη αποφάσεων σε αγορά binary option, με τη λήψη απόφασης εξαρτάται καθαρά από το αποτέλεσμα του μοντέλου. Ο αλγόριθμος που χρησιμοποιήθηκε στο πείραμα είναι ο εξής :

- Το αρχικό κεφάλαιο είναι: $btc=1$ και $usd=0$.
- Στην χρονική στιγμή i με $btc>0$ και $usd=0$, εάν $forecast[i+1] < price[i]*0,9925$
- τότε πουλάει τα btc , δηλαδή $btc=0$ και $usd= price[i]*btc*0,9925$
- Στην χρονική στιγμή i με $btc=0$ και $usd>0$, εάν $forecast[i+1] > price[i]*1,0075$, τότε αγοράζει btc , δηλαδή $btc=1$ και $usd= usd-Price[i]* 1.0075$
- Σε κάθε άλλη περίπτωση δεν πραγματοποιεί κάποια συναλλαγή.

Εικόνα 22. Διαγράμμα Ροής(Flow chart) του αλγορίθμου εφαρμογής του μοντέλου.



Υποσημείωση Διαγράμματος Ροής : Στην εκτέλεση συναλλαγών μετά τους ελέγχους , η τιμή (price[i]) συμπεριλαμβάνει το spread , δηλαδή είναι $price[i] * 1.0075$ στην αγορά του btc και $price[i] * 0.9975$ η τιμή πώλησης του

5. Παρουσίαση Αποτελεσμάτων

5.1. Αποτελέσματα επιλογής μεθόδου πρόβλεψης

Αρχικά θα παρουσιαστούν τα αποτελέσματα βάση των οποίων έγινε η επιλογή μεθόδου για την πρόοδο του μοντέλου. Πιο αναλυτικά θα παρουσιαστούν και θα συγκριθούν τα αποτελέσματα από τις μεθόδους Naïve , Holt , Ses , Arima και Moving Average. Η πρώτη διαλογή στις μεθόδους έγινε σε επίπεδο λεπτού για ορίζοντα 60 και παρατηρήσεις υλοποίησης (observations) ίσες με 45, 90, 180 , 360 και 720 . Τα αποτελέσματα αποτελούν ουσιαστικά το μέσο απόλυτο σφάλμα για ολόκληρο το δείγμα που χρησιμοποιήθηκε στη δημιουργία του μοντέλου.

5.1.1. Μέθοδος Naïve

Τα αποτελέσματα που ελήφθησαν για τη μέθοδο Naïve είναι :

Πίνακας 1. Αποτελέσματα έρευνας με τη μέθοδο Naïve

Συχνότητα (min)	1
Παρατηρήσεις	45
Άθροισμα Σφαλμάτων	7.898
Μέσο Απόλυτο Σφάλμα	9,470024
Παρατηρήσεις	90
Άθροισμα Σφαλμάτων	7.898
Μέσο Απόλυτο Σφάλμα	9,470024
Παρατηρήσεις	180
Άθροισμα Σφαλμάτων	7.898
Μέσο Απόλυτο Σφάλμα	9,470024
Παρατηρήσεις	360
Άθροισμα Σφαλμάτων	7.898
Μέσο Απόλυτο Σφάλμα	9,470024
Παρατηρήσεις	720
Άθροισμα Σφαλμάτων	7.898

Μέσο Απόλυτο Σφάλμα	9,470024
---------------------	----------

5.1.2. Μέθοδος Holt

Τα αποτελέσματα που ελήφθησαν για τη μέθοδο Holt είναι:

Πίνακας 2. Αποτελέσματα έρευνας με τη μέθοδο Holt

Συχνότητα	1
Παρατηρήσεις	45
Άθροισμα Σφαλμάτων	14.169
Μέσο Απόλυτο Σφάλμα	16,98921
Παρατηρήσεις	90
Άθροισμα Σφαλμάτων	12.794
Μέσο Απόλυτο Σφάλμα	15,34053
Παρατηρήσεις	180
Άθροισμα Σφαλμάτων	11.704
Μέσο Απόλυτο Σφάλμα	14,03357
Παρατηρήσεις	360
Άθροισμα Σφαλμάτων	10.457
Μέσο Απόλυτο Σφάλμα	12,53837
Παρατηρήσεις	720
Άθροισμα Σφαλμάτων	5.814
Μέσο Απόλυτο Σφάλμα	11,88729

5.1.3. Μέθοδος SES

Τα αποτελέσματα που ελήφθησαν για τη μέθοδο SES είναι :

Πίνακας 3. Αποτελέσματα έρευνας με τη μέθοδο Ses

Συχνότητα (min)	1
Παρατηρήσεις	45
Άθροισμα Σφαλμάτων	7.857
Μέσο Απόλυτο Σφάλμα	9,432173
Παρατηρήσεις	90
Άθροισμα Σφαλμάτων	7.856
Μέσο Απόλυτο Σφάλμα	9,430972
Παρατηρήσεις	180
Άθροισμα Σφαλμάτων	7.854
Μέσο Απόλυτο Σφάλμα	9,428571
Παρατηρήσεις	360
Άθροισμα Σφαλμάτων	7.850
Μέσο Απόλυτο Σφάλμα	9,42377
Παρατηρήσεις	720
Άθροισμα Σφαλμάτων	7.854
Μέσο Απόλυτο Σφάλμα	9,428571

5.1.4. Μέθοδος Moving Average

Τα αποτελέσματα που ελήφθησαν για τη μέθοδο Moving Average με window = 3 είναι:

Πίνακας 4. Αποτελέσματα έρευνας με τη μέθοδο Moving Average

Συχνότητα	1	1	1	1	1
Παρατηρήσεις	45	90	180	360	720
Ορίζοντας	60	60	60	60	60
Άθροισμα Σφαλμάτων	7916	10648	14257	20770	26144
Μέσο Απόλυτο Σφάλμα	9,503	12,78	17,11	24,93	31,38

5.1.5. Μέθοδος ARIMA

Τα αποτελέσματα που ελήφθησαν για τη μέθοδο ARIMA είναι:

Πίνακας 5. Αποτελέσματα έρευνας με τη μέθοδο ARIMA

Συχνότητα	1	1	1	1	1
Παρατηρήσεις	45	90	180	360	720
Ορίζοντας	60	60	60	60	60
Άθροισμα Σφαλμάτων	8144	8666	9674	11674	15804
Μέσο Απόλυτο Σφάλμα	9,7768	10,403	11,61345	14,014	18,972

5.2. Αποτελέσματα προβλέψεων χρονικής συνάθροισης

Στη συνέχεια και έχοντας επιλέξει ως μέθοδο πρόβλεψης τη SES, παρατίθενται αναλυτικά τα αποτελέσματα των προβλέψεων για κάθε ένα από τα 5 διαφορετικά επίπεδα χρονικής συνάθροισης. Λόγο του πολύ μεγάλου όγκου των δεδομένων θα παρουσιαστούν για κάθε επίπεδο χρονικής συνάθροισης ένα μέρος από τα αποτελέσματα των προβλέψεων.

Πιο συγκεκριμένα παρατίθενται προβλέψεις των 100 τελευταίων ωρών του δείγματος για τα εξής :

5.2.1. Επίπεδο συνάθροισης (aggregation) 1 λεπτού

Πίνακας 6. Αποτελέσματα επιπέδου συνάθροισης 1 λεπτού

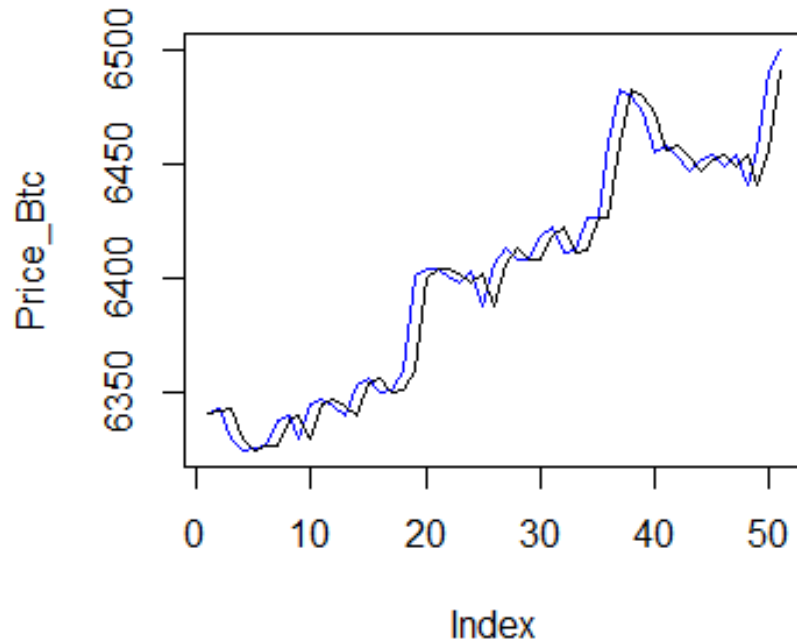
Παρατηρήσεις Υλοποίησης	120	240	360	480	600
1	6.202,9	6.202,9	6.202,9	6.202,9	6.202,9
2	6.223,0	6.223,0	6.223,0	6.223,0	6.223,0
3	6.208,2	6.208,5	6.208,5	6.208,6	6.208,6
4	6.206,7	6.206,8	6.206,9	6.207,0	6.207,0
5	6.199,1	6.199,1	6.199,1	6.199,1	6.199,1
6	6.226,7	6.226,9	6.228,9	6.228,9	6.228,9
7	6.225,7	6.225,5	6.225,3	6.225,3	6.225,3
8	6.218,7	6.218,7	6.218,7	6.218,7	6.218,7
9	6.201,7	6.201,7	6.201,7	6.201,7	6.201,7
10	6.196,9	6.196,9	6.196,9	6.196,9	6.196,9
11	6.189,6	6.189,6	6.189,6	6.189,6	6.189,6
12	6.181,7	6.181,8	6.181,8	6.181,8	6.181,8
13	6.170,3	6.170,3	6.170,3	6.170,3	6.170,3
14	6.150,2	6.150,2	6.150,2	6.150,2	6.150,2
15	6.152,8	6.152,8	6.152,8	6.152,8	6.152,8
16	6.193,6	6.193,6	6.193,6	6.193,6	6.193,7
17	6.192,0	6.192,1	6.192,0	6.192,0	6.191,7
18	6.201,8	6.201,8	6.201,8	6.201,8	6.201,8
19	6.195,6	6.196,0	6.195,3	6.195,2	6.195,0
20	6.192,5	6.192,5	6.192,5	6.192,5	6.192,5

Τα μέσα απόλυτα σφάλματα των παραπάνω προβλέψεων για ολόκληρο το δείγμα είναι αντίστοιχα:

Πίνακας 7. Μέσα απόλυτα σφάλματα για ολόκληρο το δείγμα (1 λεπτού)

Παρατηρήσεις Υλοποίησης	120	240	360	480	600
ΜΑΕ	9,41635	9,41695	9,417142	9,417142	9,417143

Συνεπώς για το επίπεδο λεπτού οι πιο ακριβείς προβλέψεις προέρχονται για 120 παρατηρήσεις υλοποίησης.



Εικόνα 23. Γραφική απεικόνιση προβλέψεων (μπλε χρώμα) / πραγματικών τιμών , 50 ωρών για το Set 1 min.

5.2.2. Επίπεδο συνάθροισης (aggregation) 5 λεπτού

Πίνακας 8. Αποτελέσματα επιπέδου συνάθροισης 5 λέπτου

Παρατηρήσεις Υλοποίησης	60	120	180	240	300
1	6.204,5	6.204,5	6.204,5	6.204,5	6.204,5
2	6.222,4	6.222,4	6.222,4	6.222,4	6.222,4
3	6.205,5	6.205,5	6.205,5	6.205,5	6.205,5
4	6.206,3	6.206,3	6.206,3	6.206,3	6.206,3
5	6.198,3	6.198,3	6.198,3	6.198,3	6.198,3
6	6.217,6	6.217,6	6.217,6	6.217,6	6.217,6
7	6.223,3	6.223,3	6.223,3	6.223,3	6.223,3
8	6.215,7	6.215,7	6.215,7	6.215,7	6.215,7
9	6.201,7	6.201,7	6.201,7	6.201,7	6.201,7
10	6.198,9	6.198,9	6.198,9	6.198,9	6.198,9
11	6.190,7	6.190,7	6.190,7	6.190,7	6.190,7
12	6.180,6	6.180,6	6.180,6	6.180,6	6.180,6
13	6.168,6	6.168,6	6.168,6	6.168,6	6.168,6
14	6.149,3	6.149,3	6.149,3	6.149,3	6.149,3

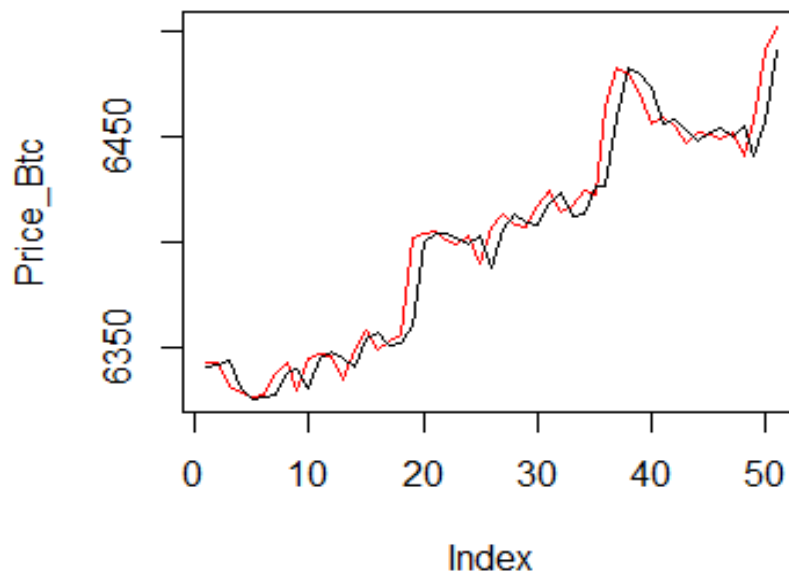
Παρατηρήσεις Υλοποίησης	60	120	180	240	300
15	6.154,9	6.154,9	6.154,9	6.154,9	6.154,9
16	6.189,5	6.189,5	6.189,5	6.189,5	6.189,5
17	6.200,5	6.200,5	6.200,5	6.200,5	6.200,5
18	6.202,7	6.202,7	6.202,7	6.202,7	6.202,7
19	6.200,3	6.200,3	6.200,3	6.200,3	6.200,3
20	6.195,2	6.195,2	6.195,2	6.195,2	6.195,2

Τα μέσα απόλυτα σφάλματα των παραπάνω προβλέψεων για ολόκληρο το δείγμα είναι αντίστοιχα:

Πίνακας 9. Μέσα απόλυτα σφάλματα για ολόκληρο το δείγμα (5 λεπτών)

Παρατηρήσεις Υλοποίησης	60	120	180	240	300
MAE	9,398631	9,390521	9,388662	9,392173	9,322622

Συνεπώς για το επίπεδο λεπτού οι πιο ακριβείς προβλέψεις προέρχονται για 180 παρατηρήσεις υλοποίησης ή $180 * 5 = 900$ λεπτά ή 15 ώρες πραγματικού χρόνου.



Εικόνα 24. Γραφική απεικόνιση προβλέψεων (κόκκινο χρώμα) / πραγματικών τιμών , 50 ωρών για το Set 5 min.

5.2.3. Επίπεδο συνάθροισης (aggregation) 15 λεπτών

Πίνακας 10. Αποτελέσματα επιπέδου συνάθροισης 15 λεπτού

Παρατηρήσεις Υλοποίησης	30	60	90	120	150	180
1	6.210,8	6.210,9	6.211,0	6.210,8	6.210,8	6.220,7
2	6.204,8	6.205,2	6.205,5	6.204,8	6.204,8	6.214,8
3	6.220,5	6.220,2	6.220,5	6.220,5	6.220,5	6.202,9
4	6.205,7	6.205,3	6.205,8	6.205,8	6.205,8	6.223,0
5	6.206,0	6.206,0	6.206,0	6.206,0	6.206,0	6.208,6
6	6.198,8	6.198,9	6.198,8	6.198,8	6.198,8	6.207,0
7	6.216,5	6.216,5	6.217,0	6.217,0	6.217,0	6.199,1
8	6.218,2	6.218,2	6.218,4	6.218,4	6.218,4	6.228,9
9	6.212,2	6.212,3	6.212,3	6.212,3	6.212,3	6.225,3
10	6.201,2	6.201,3	6.202,6	6.202,6	6.202,6	6.218,7
11	6.198,6	6.198,6	6.198,3	6.198,3	6.198,3	6.201,7
12	6.184,9	6.185,0	6.186,7	6.186,7	6.186,7	6.197,0
13	6.180,0	6.180,0	6.179,6	6.179,6	6.179,6	6.189,6
14	6.173,9	6.173,9	6.173,8	6.173,8	6.173,8	6.181,8
15	6.155,7	6.155,6	6.155,6	6.155,6	6.155,6	6.170,3
16	6.157,5	6.158,2	6.158,2	6.158,2	6.158,2	6.150,2
17	6.184,8	6.184,4	6.184,4	6.184,4	6.184,4	6.152,8
18	6.202,1	6.202,8	6.202,8	6.202,8	6.202,8	6.193,8
19	6.199,6	6.202,4	6.202,4	6.202,4	6.202,4	6.191,7
20	6.200,3	6.200,4	6.200,4	6.200,4	6.200,4	6.201,8

Τα μέσα απόλυτα σφάλματα των παραπάνω προβλέψεων για ολόκληρο το δείγμα είναι αντίστοιχα :

Πίνακας 11. Μέσα απόλυτα σφάλματα για ολόκληρο το δείγμα (15 λεπτών)

Παρατηρήσεις Υλοποίησης	30	60	90	120	150
ΜΑΕ	9,339927	9,327065	9,310126	9,319657	9,322622

Συνεπώς για το επίπεδο 15 λεπτού οι πιο ακριβείς προβλέψεις προέρχονται για 90 παρατηρήσεις υλοποίησης ή $90 * 15 = 1350$ λεπτά ή 22,5 ώρες πραγματικού χρόνου.

5.2.4. Επίπεδο συνάθροισης (aggregation) 30 λεπτών

Πίνακας 12. Αποτελέσματα επιπέδου συνάθροισης 30 λεπτού

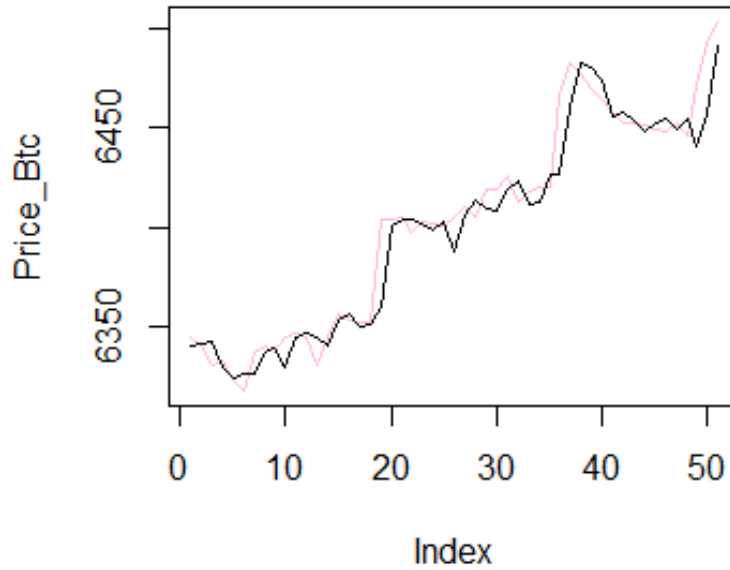
Παρατηρήσεις Υλοποίησης	15	30	45	90	180
1	6.211,9	6.211,9	6.211,9	6.211,9	6.211,9
2	6.207,4	6.207,4	6.207,4	6.207,4	6.207,4
3	6.218,9	6.218,9	6.218,9	6.218,9	6.218,9
4	6.203,3	6.203,3	6.203,3	6.203,3	6.203,3
5	6.205,7	6.205,7	6.205,7	6.205,7	6.205,7
6	6.199,0	6.199,0	6.199,0	6.199,0	6.199,0
7	6.214,5	6.214,5	6.214,5	6.214,5	6.214,5
8	6.216,9	6.216,9	6.216,9	6.216,9	6.216,9
9	6.211,8	6.211,8	6.211,8	6.211,8	6.211,8
10	6.197,3	6.197,0	6.197,3	6.197,3	6.197,3
11	6.200,1	6.199,9	6.200,1	6.200,1	6.200,1
12	6.180,1	6.179,8	6.180,1	6.180,1	6.180,1
13	6.181,8	6.181,8	6.181,8	6.181,8	6.181,8
14	6.175,1	6.175,1	6.175,1	6.175,1	6.175,1
15	6.157,6	6.157,6	6.157,6	6.157,6	6.157,6
16	6.155,2	6.153,4	6.153,4	6.153,4	6.153,4
17	6.186,2	6.185,5	6.185,5	6.185,5	6.185,5
18	6.199,2	6.199,8	6.199,8	6.199,8	6.199,8
19	6.186,5	6.186,6	6.186,6	6.186,6	6.186,6
20	6.203,9	6.199,2	6.199,2	6.199,2	6.199,2

Τα μέσα απόλυτα σφάλματα των παραπάνω προβλέψεων για ολόκληρο το δείγμα είναι αντίστοιχα :

Πίνακας 13. Μέσα απόλυτα σφάλματα για ολόκληρο το δείγμα (30 λεπτών)

Παρατηρήσεις Υλοποίησης	15	30	45	90	180
ΜΑΕ	9,50927	9,506065	9,457377	9,451787	9,440917

Συνεπώς για το επίπεδο 30 λεπτού οι πιο ακριβείς προβλέψεις προέρχονται για 180 παρατηρήσεις υλοποίησης ή $180 * 30 = 5400$ λεπτά ή 90 ώρες πραγματικού χρόνου.



Εικόνα 25. Γραφική απεικόνιση προβλέψεων (ροζ χρώμα) / πραγματικών τιμών , 50 ωρών για το Set 30 min.

5.2.5. Επίπεδο συνάθροισης (aggregation) 60 λεπτών

Πίνακας 14. Αποτελέσματα επιπέδου συνάθροισης 60 λεπτών

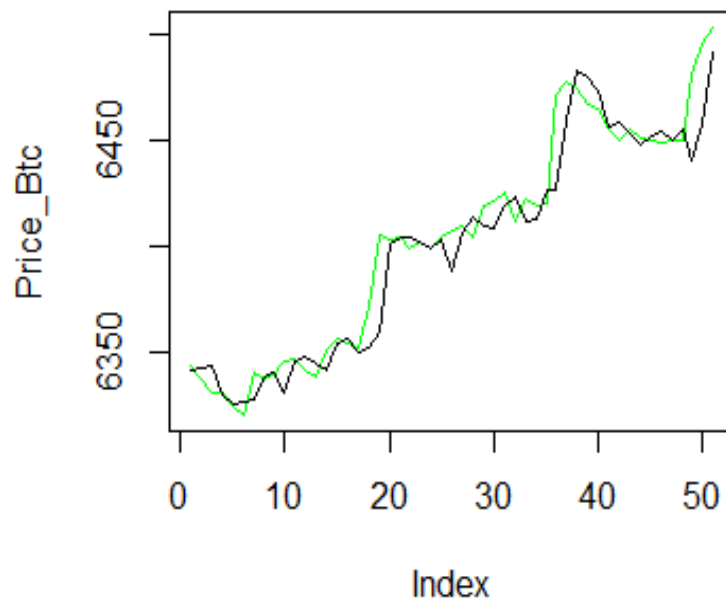
Παρατηρήσεις Υλοποίησης	12	24	36	48	60
1	6.209,1	6.209,1	6.209,1	6.209,1	6.209,1
2	6.211,8	6.212,0	6.211,8	6.211,8	6.211,8
3	6.214,9	6.214,4	6.214,9	6.214,9	6.214,9
4	6.203,2	6.203,2	6.203,1	6.203,1	6.203,1
5	6.203,9	6.203,9	6.203,9	6.203,9	6.203,9
6	6.203,9	6.204,6	6.203,9	6.203,9	6.203,9
7	6.217,3	6.217,4	6.217,3	6.217,3	6.217,3
8	6.220,0	6.219,3	6.220,0	6.220,0	6.220,0
9	6.211,6	6.209,1	6.211,6	6.211,6	6.211,6
10	6.195,1	6.194,4	6.195,1	6.195,1	6.195,1
11	6.195,9	6.195,9	6.195,9	6.195,9	6.195,9
12	6.177,2	6.177,4	6.177,2	6.177,2	6.177,2
13	6.180,0	6.172,9	6.180,0	6.180,0	6.180,0
14	6.164,2	6.165,5	6.167,8	6.167,8	6.167,8
15	6.147,4	6.167,9	6.147,4	6.147,4	6.147,4
16	6.179,2	6.168,3	6.168,3	6.168,3	6.168,3
17	6.200,7	6.188,8	6.188,8	6.188,8	6.188,8
18	6.201,9	6.197,5	6.197,5	6.197,5	6.197,5
19	6.202,7	6.184,7	6.184,7	6.184,7	6.184,7
20	6.203,1	6.197,2	6.197,2	6.197,2	6.197,2

Τα μέσα απόλυτα σφάλματα των παραπάνω προβλέψεων για ολόκληρο το δείγμα είναι αντίστοιχα :

Πίνακας 15. Μέσα απόλυτα σφάλματα για ολόκληρο το δείγμα (60 λεπτών)

Παρατηρήσεις Υλοποίησης	12	24	36	48	60
MAE	10,2800	10,1631	10,00998	10,00751	10,01933

Συνεπώς για το επίπεδο 60 λεπτού οι πιο ακριβείς προβλέψεις προέρχονται για 48 παρατηρήσεις υλοποίησης ή $48 * 60 = 2880$ λεπτά ή 48 ώρες πραγματικού χρόνου.



Εικόνα 26. Γραφική απεικόνιση προβλέψεων (πράσινο χρώμα) / πραγματικών τιμών , 50 ωρών για το Set 60 min.

5.3. Αποτελέσματα προβλέψεων με ΜΑΡΑ

Παρακάτω παρουσιάζονται τα αποτελέσματα μετά το συνδυασμό των επιμέρους επιπέδων χρονικής συνάθροισης . Αξίζει να σημειωθεί ότι η πιο ακριβής πρόβλεψη σε αυτό το στάδιο της εργασίας είναι για συχνότητα 15λεπτών (set15) και 48 παρατηρήσεις υλοποίησης (12 ώρες σε πραγματικό χρόνο) με απόλυτο μέσο σφάλμα 9.319657. Στη συνέχεια με αναρτώνται τα αποτελέσματα που υπολογίσθηκαν με την εκτέλεση της συνάρτησης optim:

$$w_1 = 0.4425617 \quad w_2 = 0.00002386825 \quad w_3 = 0.2746586$$

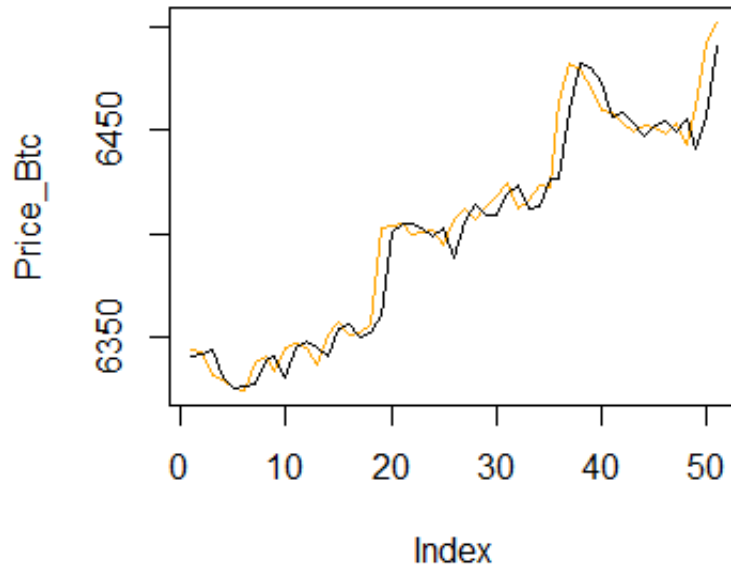
$$w_4 = 0.199256 \quad w_5 = 0.08349906$$

Με χρήση των βαρών σε κάθε επίπεδο χρονικής οι συνάθροισης το δείγμα(20 παρατηρήσεων) των αντίστοιχων τελικών προβλέψεων είναι :

Πίνακας 16. Αποτελέσματα προβλέψεων με ΜΑΡΑ

1	6.212,7
χ2	6.205,3
3	6.220,8
4	6.206,3
5	6.206,2
6	6.199,4
7	6.221,8
8	6.221,3
9	6.215,0
10	6.200,5
11	6.197,9
12	6.185,9
13	6.181,0
14	6.172,0
15	6.152,9
16	6.155,7
17	6.189,1
18	6.197,0
19	6.197,5
20	6.197,7

Το μέσο απόλυτο σφάλμα του συνδυασμού των 5 επιπέδων χρονικής συνάθροισης είναι **MAE = 9.179007**.



Εικόνα 27. Γραφική απεικόνιση προβλέψεων (πορτοκαλί χρώμα) / πραγματικών τιμών , 50 ωρών για το συνδυασμό όλων των επιπέδων χρονικής συνάθροισης.

5.4. Αποτελέσματα πειράματος εφαρμογής

Ο αλγόριθμος που περιγράφηκε στο κεφάλαιο 4 εκτελέστηκε εισάγοντας τα παρακάτω δεδομένα στο μοντέλο::

Πίνακας 17. Πίνακας δεδομένων εισαγωγής

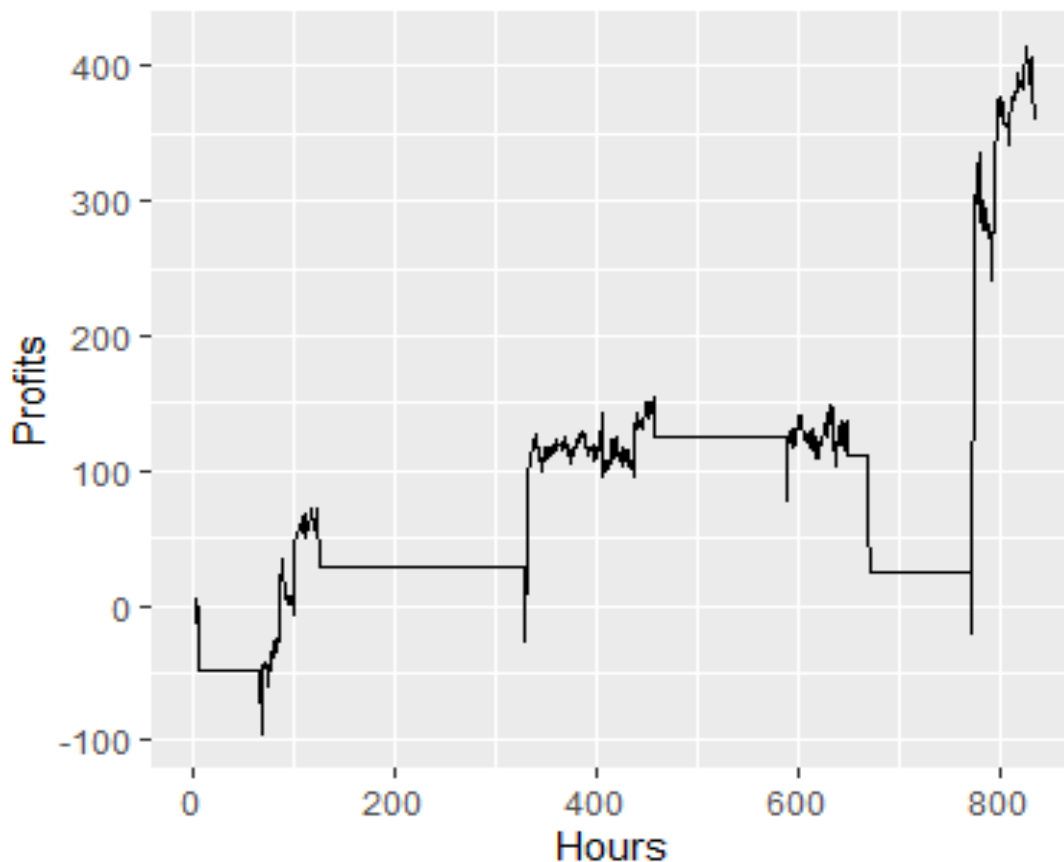
United States Dollar (\$)	6396.450
Bitcoins	0
Exchange Spread	0.75%

Το αρχικό ποσό των δολαρίων επιλέχθηκε διότι την αρχική στιγμή(πρώτη τιμή στις παρατηρήσεις του δείγματος) η τιμή του bitcoin ήταν 6897.372\$. Ο παραπάνω αλγόριθμος έτρεξε για 50000 λεπτά ή 833 ώρες , για ολόκληρο το δείγμα. Με αρχικό κεφάλαιο 6369,45 \$ και 0 btc , εκτελέστηκαν 17 συναλλαγές με έσοδα 272.13 \$. Αναλυτικά τα αποτελέσματα παρουσιάζονται στο Κεφάλαιο 5.

Πίνακας 18. Πίνακας αποτελεσμάτων

United States Dollar(\$)	6897.372
Bitcoins	0
Net Profits(\$)	359,56
Internal Rate of Return(IRR)	18,48%
Number of Iterations	833
Transactions Completed	17

Τα αποτελέσματα της χρήσης του μοντέλου είναι εξαιρετικά καθώς κέρδη 359.56 \$ ή 18.48% του αρχικού κεφαλαίου απόδοση σε μόλις 833 ώρες είναι πολύ παραπάνω από ικανοποιητική. Για την καλύτερη κατανόηση της ροής αλγορίθμου παραθέεται παρακάτω ένα διάγραμμα που ακολουθεί την πορεία το κερδών στην εξέλιξη του χρόνου και διαφαίνεται σε ποιες χρονικές περιόδους εκτελέσθηκαν συναλλαγές , ποιες ήταν επικερδείς συναλλαγές και ποιες ζημιογόνες.



Εικόνα 28. Απεικόνιση των κερδών στην εξέλιξη του χρόνου.

Όπως φαίνεται και στο διάγραμμα αρχικά πραγματοποιούνται κάποιες συναλλαγές που ζημιώνουν προσωρινά τον επενδυτή και για τις πρώτες 75 περίπου ώρες δεν υλοποιούνται άλλες συναλλαγές. Έπειτα για τις επόμενες 50 ώρες πυροδοτούνται σχεδόν κάθε ώρα συναλλαγές , αρχικά ζημιογόνες αλλά τελικά επικερδής με αποτέλεσμα μετά τις πρώτες 120 ώρες ή 5 ωρολογιακές μέρες ο , το μοντέλο να έχει αποδώσει περίπου 75 \$ ή 1.2 % του αρχικού κεφαλαίου. Τις επόμενες 200 ώρες δεν πραγματοποιείται κάποια συναλλαγή. Αμέσως μετά, παρατηρείται μία αρκετά κακή επιλογή συναλλαγής (μείον 50 \$) που όμως αμέσως διορθώνεται με άνοδο κοντά στα 150 \$ κέρδη που συντηρείται για 350 ώρες ακόμα. Τέλος, ακολουθεί μία ακόμη αρνητική επιλογή συναλλαγής με απώλειες 80 \$, η οποία γρήγορα θα υπερκαλυφθεί από συνεχόμενες συναλλαγές 50 ωρών που θα εκτοξεύσουν τα τελικά κέρδη σε 359,56 \$ ή 18,48% των ίδιων αρχικών κεφαλαίων.

6. Συμπεράσματα και προοπτικές περαιτέρω έρευνας

6.1. Συμπεράσματα

Το βασικό ζήτημα το οποίο επιχειρήθηκε να απαντηθεί μέσω της εκπόνησης της παρούσας διπλωματικής εργασίας είναι το κατά πόσο στατιστικές μέθοδοι μπορούν χρησιμοποιηθούν ως κύριο εργαλείο για την πρόβλεψη κερδοφόρων συναλλαγών κρυπτονομισμάτων σε ωριαία βάση. Από τα αποτελέσματα που παρουσιάστηκαν στα προηγούμενα δύο κεφάλαια είναι εμφανές πως η εφαρμογή στατιστικών μεθόδων προβλέψεων είναι ικανή να οδηγήσει σε αποδόσεις ακόμα και 19% σε 833 ώρες ή περίπου 14 μέρες. Η βασική ιδέα είναι ότι ένας επενδυτής ή trader, έχοντας στη διάθεσή του το εργαλείο που δημιουργήθηκε και δεδομένα των προηγούμενων 2 ημερών μπορεί να κάνει προβλέψεις για την επόμενη ώρα, να εκτελεί τις συναλλαγές που του υποδεικνύει το πρόγραμμα και στη συνέχεια να ανανεώνει τα δεδομένα του με την παρούσα τιμή του κρυπτονομίσματος.

Η παρούσα μελέτη έδειξε πως μια απλούστερη προσέγγιση του πολύπλοκου προβλήματος της πρόβλεψης της τιμής κρυπτονομισμάτων μπορεί να είναι αρκετά αποδοτική. Συνήθως προτείνεται η πρόβλεψη τιμών με μεγάλες διακυμάνσεις με πολυσύνθετα νευρωνικά δίκτυα που απαιτούν τεράστιο όγκο δεδομένων για να εκπαιδευτούν και αμφίβολα αποτελέσματα. Κατά την άποψη μου, ο trader που εκτελεί τόσο συχνές συναλλαγές χρειάζεται ευελιξία και ταχύτητα στο εργαλείο που χρησιμοποιεί. Παρόλα αυτά, χρειάζεται ιδιαίτερη προσοχή στην διαχείριση των αποτελεσμάτων αυτών. Πιθανή εφαρμογή των βέλτιστων μεθόδων στην αγορά δεν συνεπάγεται άμεσα ίδια κέρδη με αυτά που υπολογίστηκαν στο πείραμα αυτό. Πέρα των υποθέσεων και των απλοποιήσεων που πραγματοποιήθηκαν στην προετοιμασία των δεδομένων για ευκολία περάτωσης του πειράματος, δεν πρέπει να θεωρηθεί αμελητέο ότι τα δεδομένα και κατ' επέκταση και το μοντέλο που δημιουργήθηκε αφορούν μια συγκεκριμένη χρονική περίοδο, όπου η δυναμική ψυχολογία της αγοράς είχε ορισμένα χαρακτηριστικά. Επομένως, η χρήση του μοντέλου σε αρκετά μεταγενέστερη χρονική περίοδο με διαφορετικά χαρακτηριστικά δεν μπορεί να θεωρηθεί δεδομένη.

6.2. Μελλοντικές Προεκτάσεις

6.2.1. Γενικά

Στο σημείο αυτό, δίνεται η δυνατότητα να παρουσιαστούν ορισμένα εγχειρήματα τα οποία διατηρήθηκαν εκτός του φάσματος εξέτασης της συγκεκριμένης διπλωματικής εργασίας, αλλά η μελέτη και η ανάπτυξή τους μπορεί να προσδώσει ένα πιο ολοκληρωμένο επενδυτικό εργαλείο.

6.2.2. Δημιουργία εργαλείου που εκτελεί τις συναλλαγές

Στα πλαίσια δημιουργίας ενός καθολικού επενδυτικού πακέτου , που μειώνει την πιθανότητα λάθους από το χρήστη αλλά παράλληλα βελτιώνει την απόδοση και την ευρωστία (robustness) του εργαλείου με την αυτοματοποίηση των διαδικασιών , προτείνεται η δημιουργία ενός προγράμματος εκτέλεσης συναλλαγών. Αναλυτικότερα, πρόκειται για μία εφαρμογή με ένα περιβάλλον στο οποίο θα περιηγείται ο trader και το οποίο θα παίρνει χειροκίνητα την παρούσα τιμή του κρυπτονομίσματος. Έχοντας ενσωματωμένο το μοντέλο που δημιουργήθηκε στην παρούσα εργασία θα λαμβάνει σαν είσοδο το αποτέλεσμα αν θα πρέπει να εκτελέσει συναλλαγή και ποια θα είναι αυτή. Το πρόγραμμα θα εκτελεί μόνο του τη συναλλαγή μέσω του συναλλακτηρίου. Τα μεγαλύτερα συναλλακτήρια επιτρέπουν τέτοιου είδους αυτοματοποιημένες συναλλαγές, με επικοινωνία μέσω API. Για παράδειγμα επισυνάπτεται η σελίδα με τις εντολές του exchange HuobiPro, (<https://huobiapi.github.io/docs/spot/v1/en/#place-a-new-order>)

6.2.3. Δημιουργία εργαλείου που συλλέγει τα δεδομένα

Επιπρόσθετα, επεκτείνοντας την προηγούμενη πρόταση , σε συνδυασμό με την υπάρχουσα μέθοδο , προτείνεται η κατασκευή προγράμματος επικοινωνίας με το συναλλακτήριο μέσω API για τη συλλογή δεδομένων σε πραγματικό χρόνο. Δηλαδή, ο trader να λαμβάνει κάθε αλλαγή της τιμής του κρυπτονομίσματος , την ίδια χρονική στιγμή που αλλάζει και στο συναλλακτήριο. Αυτό θα ολοκληρώσει ουσιαστικά την όλη διαδικασία , καθώς τόσο τα δεδομένα όσο και η λήψη απόφασης συναλλαγής αλλά και η εκτέλεση αυτής , πραγματοποιούνται από το ίδιο το πρόγραμμα. Κατ' επέκταση, εφόσον υπάρχει η δυνατότητα απόκτησης των δεδομένων σε πραγματικό χρόνο, το παρών πρόγραμμα μπορεί με παραμετροποίηση να λαμβάνει αποφάσεις για πολύ πιο σύντομες συναλλαγές, δημιουργώντας ουσιαστικά ένα ολοκληρωμένο εργαλείο High-Frequency-Trading (HFT).

6.2.4. Εναλλακτικές μέθοδοι πρόβλεψης

Σε ότι αφορά την επίδοση και την ιδιοσυγκρασία του μοντέλου που δημιουργήθηκε στο πλαίσιο εκπόνησης αυτής της Διπλωματικής, μία πρόταση που ενδεχομένως να βελτιώσει την απόδοση των προβλέψεων είναι η δοκιμή περισσότερων μεθόδων πρόβλεψης. Πιο συγκεκριμένα, θα ήταν εύλογο να δοκιμαστούν τόσο μοντέλα γραμμικής παλινδρόμησης στα διάφορα επίπεδα χρονικής

συνάθροισης αλλά και η μέθοδος Theta, ώστε να δοθεί μία πιο σφαιρική άποψη για τις προβλέψεις της δοθείσας χρονοσειράς. Πιθανόν τα αποτελέσματα να είναι τέτοια ώστε να χρησιμοποιηθούν διαφορετικές μέθοδοι ανά επίπεδο και ύστερα να συνδυαστούν για την τελική πρόβλεψη. Επίσης, δεδομένου ότι υπάρχει αρκετά μεγάλο πλήθος δεδομένων , προτείνεται η διερεύνηση της αρχιτεκτονικής και των υπερπαραμέτρων ενός νευρωνικού δικτύου που με γνώμονα την ελαχιστοποίηση του σφάλματος , θα έχει καλύτερα αποτελέσματα.

Βιβλιογραφία

1. Πετρόπουλος, Φ. και Ασημακόπουλος, Β. (2011), *Επιχειρησιακές Προβλέψεις*, Αθήνα: Εκδόσεις Συμμετρία.
2. Hotta, L. K.; Morettin, P.A. and Valls Pereira, P.L. (1992). *The Effect of Overlapping Aggregation on Time Series Models: An Application to the Unemployment Rate in Brazil*. *Brazilian Review of Econometrics*, Vol. 12, No. 2.
3. Hyndman, R.J. and Athanasopoulos, G. (2018). *Forecasting: Principles and Practice*. OTexts.
4. Kourentzes, N. and Petropoulos, F. (2014). *Improving forecasting via multiple temporal aggregation*. In: *Foresight: The International Journal of Applied Forecasting*, Vol. 2014, No. 34, 2014, p. 12-17.
5. Makridakis, S.G.; Wheelwright, S.C. and Hyndman, R.J. (1998). *Forecasting: Methods and Applications – 3rd Edition*, John Wiley and Sons.
6. Petropoulos, F.; Kourentzes, N. and Nikolopoulos, K.. (2016). *Another look at estimators for intermittent demand*. *International Journal of Production Economics*. Volume 181, Part A, November 2016, Pages 154-161.
7. *Bitcoin: A Peer-to-Peer Electronic Cash System*, Satoshi Nakamoto, October 2008
8. *An Application Of Time Series Arima Forecasting Model For Predicting Sugarcane Production In India*. Kumar, Manoj & Anand, Madhu. (2014)
9. A Study on Exponential Smoothing Method for Forecasting. Shastri, Sourabh & Sharma, Amardeep & Mansotra, Vibhakar & Sharma, Anand & Bhadwal, Arun & Kumari, Monika. (2018). *International Journal of Computer Sciences and Engineering*.
10. Numerical Optimization in R : Beyond optim. Varadhan, Ravi. (2014). *Journal of Statistical Software*.
11. Bitcoin Pricing, Adoption, and Usage: Theory and Evidence, Susan Athey, Ivo Parashkevov ,Vishnu Sarukkai, Jing Xia. August,2016. Stanford University

Ηλεκτρονικές Πηγές

1. Bitcoin (2019). Διαθέσιμο διαδικτυακά : <https://bitcoin.org/el/how-it-works> , <https://nakamotoinstitute.org/static/docs/bitcoin.pdf>, <https://digiconomist.net/bitcoin-energy-consumption>.
2. Democracy Earth (2019). Διαθέσιμο διαδικτυακά: <https://democracy.earth/>.
3. Dubai Digital Strategy (2019). Διαθέσιμο διαδικτυακά: <https://www.smartdubai.ae/initiatives/blockchain>
4. Ethereum (2019). Διαθέσιμο διαδικτυακά: <https://ethereum.org/>.
5. Energy Web Foundation (2019). Διαθέσιμο διαδικτυακά: <https://www.energyweb.org/about/what-we-do/>.
6. Follow My Vote (2019). Διαθέσιμο διαδικτυακά: <https://followmyvote.com/>.
7. Highcharts (2019). Διαθέσιμο διαδικτυακά: <https://www.highcharts.com>.
8. Huobi Exchange (2019). Διαθέσιμο διαδικτυακά: <https://huobiapi.github.io/docs/spot/v1/en/#place-a-new-order>.
9. Investopedia (2019). Διαθέσιμο διαδικτυακά: <https://www.investopedia.com/articles/active-trading/101014/basics-algorithmic-trading-concepts-and-examples.asp>.
10. Multichain, Enterprise Blockchain (2019). Διαθέσιμο διαδικτυακά: <https://www.multichain.com/>.
11. New York Stock Exchange (2019). Διαθέσιμο διαδικτυακά: <https://www.nyse.com/markets/liquidity-programs>.
12. New York Digital Investment Group (2019). Διαθέσιμο διαδικτυακά: <https://nydig.com/>.
13. OkeX Exchange (2019). Διαθέσιμο διαδικτυακά: <https://okexsupport.zendesk.com/hc/en-us/articles/360027486392-Spot-Trading-What-is-Time-weighted-average-price-TWAP->.
14. OpenBazaar (2019). Διαθέσιμο διαδικτυακά: <https://openbazaar.org/>.
15. Stampd, ελληνική blockchain start-up. (2019) Διαθέσιμο διαδικτυακά: <https://stampd.io>.
16. Ubitquity (2019) Διαθέσιμο διαδικτυακά: (<https://www.ubitquity.io/>)

Παραρτήματα

Παράρτημα κώδικα R

Π1. Βιβλιοθήκες

```
library(readxl)
library("zoo")
library("dlsem", lib.loc=~R/win-library/3.5")
library("TestDataImputation", lib.loc=~R/win-library/3.5")
library("forecast", lib.loc=~R/win-library/3.5")
library("fpp2", lib.loc=~R/win-library/3.5")
library("imputeTS", lib.loc=~R/win-library/3.5")
library(lubridate)
library(forecast)
library(smooth)
require(smooth)
require(greybox)
require(Mcomp)
library(openxlsx)
```

Π2. Εισαγωγή Δεδομένων

```
Price_Btc_Sept_1min<-
read_excel("C:/Users/leftferis/OneDrive/Documents/Price_Btc_Sept_1min.xlsx")
s
Price_Btc_Oct_1min<-
read_excel("C:/Users/mts2wks18/Downloads/Price_Btc_Oct_1min.xlsx")
Price_Btc_dec_1min<-
read_excel("C:/Users/mts2wks18/Downloads/Price_Btc_dec_1min.xlsx")
Price_Btc_Nov_1min<-
read_excel("C:/Users/mts2wks18/Downloads/Price_Btc_Nov_1min.xlsx")
```

```

Price_Btc_Augoust_1min<-
read_excel("C:/Users/mts2wks18/Downloads/Price_Btc_Augoust_1min.xlsx")

Price_Btc_Sept_1min<-
read_excel("C:/Users/mts2wks18/Downloads/Price_Btc_Sept_1min.xlsx")

filld<-as.data.frame(format(seq(from=ISOdate(2018,7,31,hour=0,min=0,tz="GMT"
),to=ISOdate(2018,12,18,hour = 23,min = 59,tz="GMT"),by="1 min")
))

a_set=bind.data.frame(Price_Btc_Augoust_1min,Price_Btc_Sept_1min,Price_Btc_
Oct_1min,Price_Btc_Nov_1min ,Price_Btc_dec_1min )

a_set$Timestamp<-filld

```

Π3. Μέθοδοι Πρόβλεψης

Διαχείριση ασυνήθιστων τιμών

```

dat = as.numeric(a_set$`Weighted Price`)
dat[dat==0] <- NA
dat[dat<2000]<-NA
dat[dat>9000]<-NA
dat <- na.kalman(as.numeric(dat))
a_set$`Weighted Price`<-dat

```

Δημιουργία 4 επιπέδων χρονικής συνάρτησης

```

j=2
i=6
set5[1]<-dat[1]
while (i < 202000) {
k=1
sum = 0
while(k<4){
sum <- sum + dat[i+k] + dat[i-k]
k<-k+1
}
}

```

```

set5[j]<-(sum +dat[i])/5
i<-i+5
j<-j+1
}
j=2
i=16
set15[1]<-dat[1]
while (i < 202000) {
k=1
sum = 0
while(k<8){
sum <- sum + dat[i+k] + dat[i-k]
k<-k+1
}
set15[j]<-(sum +dat[i])/15
i<-i+15
j<-j+1
}

j=2
i=31
set30[1]<-dat[1]
while (i < 202000) {
k=1
sum = 0
while(k<16){
sum <- sum + dat[i+k] + dat[i-k]
k<-k+1
}
set30[j]<-(sum )/30

```

```

i<-i+30
j<-j+1
}
j=2
i=61
set60[1]<-dat[1]
while (i < 202000) {
k=1
sum = 0
while(k<31){
sum <- sum + dat[i+k] + dat[i-k]
k<-k+1
}
set60[j]<-(sum )/60
i<-i+60
j<-j+1
}

```

Παραγωγή Προβλέψεων

```

n1<-matrix(data=0,nrow = 834 ,ncol = 6)
n15<-matrix(data=0,nrow = 834 ,ncol = 6)
n30<-matrix(data=0,nrow = 834 ,ncol = 6)
n5<-matrix(data=0,nrow = 834 ,ncol = 6)
n60<-matrix(data=0,nrow = 834 ,ncol = 6)
i=150060
j=1
while (i>100060){
n1[j,6]<-dat[i]
j<-j+1
i<-i-60
}

```



```

}
i=150060
j=1
while (i>100060){
  n15[j,6]<-dat[i]
  j<-j+1
  i<-i-60
}
i=150060
j=1
while (i>100060){
  n5[j,6]<-dat[i]
  j<-j+1
  i<-i-60
}
i=150060
j=1
while (i>100060){
  n30[j,6]<-dat[i]
  j<-j+1
  i<-i-60
}
i=150060
j=1
while (i>100060){
  n60[j,6]<-dat[i]
  j<-j+1
  i<-i-60
}
j=1

```

```

a= 5000
i=1
while (a> 3333){
  t=1
  obs = 12*2
  while (t < 6) {
    a1 = a
    a2<- a1 - (obs * t)
    k<- ts(set30[(a2:a1)])
    n30[i,t] = naive(k,h=1)$mean
    t=t+1
  }
  i=i+1
  a=a-2
}
a=2500
i=1
while (a> 1666){
  t=1
  obs = 12
  while (t < 6) {
    a1 = a
    a2<- a1 - (obs * t)
    k<- ts(set60[(a2:a1)])
    n60[i,t] = naive(k,h=1)$mean
    t=t+1
  }
  i=i+1
  a=a-1
}

```

```

a=10000
i=1
while (a> 6666){
  t=1
  obs = 12*4
  while (t < 6) {
    a1 = a
    a2<- a1 - (obs * t)
    k<- ts(set15[(a2:a1)])
    n15[i,t] = naive(k,h=1)$mean
    t=t+1
  }
  i=i+1
  a=a-4
}
a=30000
i=1
while (a> 20000){
  t=1
  obs = 12*4*2
  while (t < 6) {
    a1 = a
    a2<- a1 - (obs * t)
    k<- ts(set5[(a2:a1)])
    n5[i,t] = naive(k,h=1)$mean
    t=t+1
  }
  i=i+1
  a=a-12
}

```

```

a=150000
i=1
while (a > 100000){
  t=1
  obs = 12*4*2*2
  while (t < 6) {
    a1 = a
    a2<- a1 - (obs * t)
    k<- ts(dat[(a2:a1)])
    n1[i,t] = naive(k,h=1)$mean
    t=t+1
  }
  i=i+1
  a=a-60
}
resn<-matrix(data=0, nrow=4, ncol = 5)
a= 150060
i=1
while (i<5){
  resn[i,1]<- mean(abs(n1[,i]-n1[,6]))
  resn[i,2]<- mean(abs(n5[,i]-n5[,6]))
  resn[i,3]<- mean(abs(n15[,i]-n15[,6]))
  resn[i,4]<- mean(abs(n30[,i]-n30[,6]))
  resn[i,5]<- mean(abs(n60[,i]-n60[,6]))
  i=i+1
}
g1<-matrix(data=0,nrow = 834 ,ncol = 6)
g15<-matrix(data=0,nrow = 834 ,ncol = 6)
g30<-matrix(data=0,nrow = 834 ,ncol = 6)
g5<-matrix(data=0,nrow = 834 ,ncol = 6)

```

```

g60<-matrix(data=0,nrow = 834 ,ncol = 6)
i=150060
j=1
while (i>100060){
  g1[j,5]<-dat[i]
  j<-j+1
  i<-i-60
}
i=150060
j=1
while (i>100060){
  g15[j,5]<-dat[i]
  j<-j+1
  i<-i-60
}
i=150060
j=1
while (i>100060){
  g5[j,5]<-dat[i]
  j<-j+1
  i<-i-60
}
i=150060
j=1
while (i>100060){
  g30[j,5]<-dat[i]
  j<-j+1
  i<-i-60
}
i=150060

```

```

j=1
while (i>100060){
  g60[j,5]<-dat[i]
  j<-j+1
  i<-i-60
}
i=150060
j=1
while (i>100060){
  g1[j,5]<-dat[i]
  j<-j+1
  i<-i-60
}
j=1
a= 150000
i=1
while (a> 100000){
  a1=a
  obs = 12*2*2*2*2*2
  a2<- a1- obs
  k<- ts(dat[a2:a1])
  g1[j,1] = holt(k,h=60)$mean[60] + g1[j,1]

  obs = 12 *2*2*2*2*2
  a2<- a1- obs
  k<- ts(dat[a2:a1])
  g1[j,2] = holt(k,h=60)$mean[60]
  obs = 12*4*2*2*2*2

```

```

a2<- a1- obs
k<- ts(dat[a2:a1])
g1[j,3] = holt(k,h=60)$mean[60]
obs = 12*8*2*2*2*2
a2<- a1- obs
k<- ts(dat[a2:a1])
g1[j,4] = holt(k,h=60)$mean[60]
j=j+1
a<- a-60
}
i=1
resg=matrix(data=0,nrow = 4,ncol = 5)
while (i<5){
resg[i,1]<- mean(abs(g1[,i]-g1[,5]))
resg[i,2]<- mean(abs(g5[,i]-g5[,5]))
resg[i,3]<- mean(abs(g15[,i]-g15[,5]))
resg[i,4]<- mean(abs(g30[,i]-g30[,5]))
resg[i,5]<- mean(abs(g60[,i]-g60[,5]))
i=i+1
}
f1<-matrix(data=0,nrow = 834 ,ncol = 6)
f15<-matrix(data=0,nrow = 834 ,ncol = 6)
f30<-matrix(data=0,nrow = 834 ,ncol = 6)
f5<-matrix(data=0,nrow = 834 ,ncol = 6)
f60<-matrix(data=0,nrow = 834 ,ncol = 6)
i=150060
j=1
while (i>100060){
f1[j,6]<-dat[i]
j<-j+1

```

```

i<-i-60
}
i=150060
j=1
while (i>100060){
  f15[j,6]<-dat[i]
  j<-j+1
  i<-i-60
}
i=150060
j=1
while (i>100060){
  f5[j,6]<-dat[i]
  j<-j+1
  i<-i-60
}
i=150060
j=1
while (i>100060){
  f30[j,6]<-dat[i]
  j<-j+1
  i<-i-60
}
i=150060
j=1
while (i>100060){
  f60[j,6]<-dat[i]
  j<-j+1
  i<-i-60
}

```



```

j=1
a= 5000
i=1
while (a> 3333){
  t=1
  obs = 12*2
  while (t < 6) {

    a1 = a
    a2<- a1 - (obs )
    k<- ts(set30[(a2:a1)])
    f30[i,t] = ses(k,h=1)$mean
    t=t+1
    obs=obs*2
  }
  i=i+1
  a=a-2
}

a=2500
i=1
while (a> 1666){
  t=1
  obs = 12
  while (t < 6) {

    a1 = a
    a2<- a1 - (obs )
    k<- ts(set60[(a2:a1)])
    f60[i,t] = ses(k,h=1)$mean

```

```

obs=obs*2
  t=t+1
}
i=i+1
a=a-1
}
a=10000
i=1
while (a> 6666){

  t=1
  obs = 12
  while (t < 6) {

    a1 = a
    a2<- a1 - (obs )
    k<- ts(set15[(a2:a1)])
    f15[i,t] = ses(k,h=1)$mean
    obs=obs*2
    t=t+1
  }
  i=i+1
  a=a-4
}
a=30000
i=1
while (a> 20000){
  t=1
  obs = 12*2

```

```

while (t < 6) {

a1 = a
a2<- a1 - (obs )
k<- ts(set5[(a2:a1)])
f5[i,t] = ses(k,h=1)$mean
t=t+1
obs=obs*2
}
i=i+1
a=a-12}
a=150000
i=1
while (a> 100000){
t=1
obs = 12*4*2*2*2
while (t < 6) {

a1 = a
a2<- a1 - (obs )
k<- ts(dat[(a2:a1)])
f1[i,t] = ses(k,h=60)$mean[60]
t=t+1
obs= obs * 2
}
i=i+1
a=a-60
}
ress<-matrix(data=0, nrow=5, ncol = 5)
a= 150060

```

```

i=1
while (i<6){
  ress[i,1]<- mean(abs(f1[,5]-f1[,6]))
  ress[i,2]<- mean(abs(f5[,i]-f5[,6]))
  ress[i,3]<- mean(abs(f15[,i]-f15[,6]))
  ress[i,4]<- mean(abs(f30[,i]-f30[,6]))
  ress[i,5]<- mean(abs(f60[,i]-f60[,6]))
  i=i+1
}
ar1<-matrix(data=0,nrow = 834 ,ncol = 6)
ar15<-matrix(data=0,nrow = 834 ,ncol = 6)
ar30<-matrix(data=0,nrow = 834 ,ncol = 6)
ar5<-matrix(data=0,nrow = 834 ,ncol = 6)
ar60<-matrix(data=0,nrow = 834 ,ncol = 6)

```

```

i=150060
j=1
while (i>100060){
  ar1[j,6]<-dat[i]
  j<-j+1
  i<-i-60
}
i=150060
j=1
while (i>100060){
  ar15[j,6]<-dat[i]
  j<-j+1
  i<-i-60
}
i=150060

```

```
j=1
while (i>100060){
  ar5[j,6]<-dat[i]
  j<-j+1
  i<-i-60
}
i=150060
j=1
while (i>100060){
  ar30[j,6]<-dat[i]
  j<-j+1
  i<-i-60
}
i=150060
j=1
while (i>100060){
  ar60[j,6]<-dat[i]
  j<-j+1
  i<-i-60
}
i=150060
j=1
while (i>100060){
  ar1[j,6]<-dat[i]
  j<-j+1
  i<-i-60
}

j=1
a= 5000
```

```

i=1
while (a> 3333){
  t=1
  obs = 12*2
  while (t < 6) {
    a1 = a
    a2<- a1 - (obs * t)
    k<- ts(set30[(a2:a1)])
    ar30[i,t] = forecast(arima(k),h=2)$mean[2]
    t=t+1 }
  i=i+1
  a=a-2
}
a=2500
i=1

```

```

while (a> 1666){

  t=1

  obs = 12
  while (t < 6) {

    a1 = a
    a2<- a1 - (obs * t)
    k<- ts(set60[(a2:a1)])
    ar60[i,t] = forecast(arima(k),h=1)$mean[1]
    t=t+1
  }
  i=i+1

```

```

a=a-1
}
a=10000
i=1
while (a> 6666){
t=1
obs = 12*4
while (t < 6) {

a1 = a
a2<- a1 - (obs * t)
k<- ts(set15[(a2:a1)])
ar15[i,t] = forecast(arima(k),h=4)$mean[4]
t=t+1
}
i=i+1
a=a-4
}
a=30000
i=1
while (a> 20000){
t=1
obs = 12*4*2
while (t < 6) {

a1 = a
a2<- a1 - (obs * t)
k<- ts(set5[(a2:a1)])
ar5[i,t] = forecast(arima(k),h=12)$mean[12]
t=t+1
}
}
}

```

```

}
i=i+1
a=a-12
}
a=150000
i=1
while (a > 100000){
t=1
obs = 12*4*2*2
while (t < 6) {

a1 = a
a2<- a1 - (obs * t)
k<- ts(dat[(a2:a1)])
ar1[i,t] = forecast(arima(k),h=60)$mean[60]
t=t+1
}
i=i+1
a=a-60
}
resar<-matrix(data=0, nrow=4, ncol = 5)
a= 150060
i=1
while (i<5){
resar[i,1]<- mean(abs(ar1[,i]-ar1[,6]))
resar[i,2]<- mean(abs(ar5[,i]-ar5[,6]))
resar[i,3]<- mean(abs(ar15[,i]-ar15[,6]))
resar[i,4]<- mean(abs(ar30[,i]-ar30[,6]))
resar[i,5]<- mean(abs(ar60[,i]-ar60[,6]))
i=i+1
}

```



```

}
Εύρεση βέλτιστων βαρών
r<-matrix(data =(0) ,nrow=5)
ff1<-function(x) {
  x1 <- abs(sin(x[1]))
  x5 <- abs(sin(x[2]))
  x15 <- abs(sin(x[3]))
  x30<- abs(sin(x[4]))
  x60<-1-x1-x5-x15-x30
  mean(abs( (x1*f1[,3]+x5*f5[,4]+x15*f15[,3]+x30*f30[,3]+x60*f60[,4]) - f5[,6]))
}
sigmoid <- function(x) exp(x) / ( 1 + exp(x) )
convert <- function(x) {
  q1 <- sigmoid(x[1])
  q2 <- (1-q1) * sigmoid(x[2])
  q3 <- (1-q1-q2) * sigmoid(x[3])
  q4 <- (1-q1-q2-q3) * sigmoid(x[4])
  q5<- 1-q1-q2-q3-q4
  c(q1,q2,q3,q4,q5)
}
g <- function(x) ff1(convert(x))
p <- optim(c(0,0,0,0,0), g)
p$par<-convert(p$par)

```

Εφαρμογή μοντέλου

```

fin<-matrix(data=0,nrow = 834 ,ncol = 3)
inv<-matrix(data=0,nrow = 834 ,ncol = 3)
fin[,1]<-
p$par[1]*f1[,3]+p$par[2]*f5[,4]+p$par[3]*f15[,3]+p$par[4]*f30[,3]+p$par[5]*f60[,4]

inv[,1]<- fin[,1]-f5[,6]

```

```

sum[inv[,1]]
fin[1:10,1]
f5[1:10,6]
btc=matrix(data=0,ncol = 833)
usd=matrix(data=0,ncol = 833)
usd=0
btc=1
profits=6369.450
s=0
t=2
while(t<=833){
  if(((fin[t+1,1]*0.9925)>((1.0075)*f5[t,6])) && (btc==0)){
    s=s+1
    btc=1
    usd=usd-f5[t,6]-f5[t,6]}

  if(((fin[t+1,1]*1.0075)<(0.9925*f5[t,6])) && (btc==1)){
    btc=0
    s=s+1
    usd=usd+f5[t,6]-f5[t,6] }
  t=t+1
}
IRR= (usd-6369.5)^(1/34.7)-1
NET_PROFIT= usd-f5[833,6]

```