



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ Μ/Υ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
ΣΧΟΛΗ ΝΑΥΤΙΛΙΑΣ ΚΑΙ ΒΙΟΜΗΧΑΝΙΑΣ
ΤΜΗΜΑΤΟΣ ΒΙΟΜΗΧΑΝΙΚΗΣ ΔΙΟΙΚΗΣΗΣ & ΤΕΧΝΟΛΟΓΙΑΣ
ΔΙΑΠΑΝΕΠΙΣΤΗΜΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
«ΤΕΧΝΟ-ΟΙΚΟΝΟΜΙΚΑ ΣΥΣΤΗΜΑΤΑ»



ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΠΡΟΒΛΕΨΗ ΖΗΤΗΣΗΣ ΓΙΑ ΤΟΝ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟ ΠΑΡΑΓΩΓΗΣ ΣΕ ΒΙΟΜΗΧΑΝΙΑ ΤΡΟΦΙΜΩΝ

ΑΛΕΞΙΑ ΠΟΛΙΤΑΚΗ

ΕΠΙΒΛΕΠΩΝ

Βασίλειος Ασημακόπουλος, Καθηγητής Ε.Μ.Π.

ΦΕΒΡΟΥΑΡΙΟΣ 2021



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ Μ/Υ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
ΣΧΟΛΗ ΝΑΥΤΙΛΙΑΣ ΚΑΙ ΒΙΟΜΗΧΑΝΙΑΣ
ΤΜΗΜΑΤΟΣ ΒΙΟΜΗΧΑΝΙΚΗΣ ΔΙΟΙΚΗΣΗΣ & ΤΕΧΝΟΛΟΓΙΑΣ
ΔΙΑΠΑΝΕΠΙΣΤΗΜΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
«ΤΕΧΝΟ-ΟΙΚΟΝΟΜΙΚΑ ΣΥΣΤΗΜΑΤΑ»



ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΠΡΟΒΛΕΨΗ ΖΗΤΗΣΗΣ ΓΙΑ ΤΟΝ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟ ΠΑΡΑΓΩΓΗΣ ΣΕ ΒΙΟΜΗΧΑΝΙΑ ΤΡΟΦΙΜΩΝ

ΑΛΕΞΙΑ ΠΟΛΙΤΑΚΗ

ΕΠΙΒΛΕΠΩΝ

Βασίλειος Ασημακόπουλος, Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την x^η Φεβρουαρίου 2021.

(Υπογραφή)

.....

Βασίλειος Ασημακόπουλος
Καθηγητής Ε.Μ.Π.

(Υπογραφή)

.....

Ιωάννης Ψαρράς
Καθηγητής Ε.Μ.Π.

(Υπογραφή)

.....

Χρυσόστομος Δούκας
Αναπλ. Καθηγητής Ε.Μ.Π.

ΦΕΒΡΟΥΑΡΙΟΣ 2021

(Υπογραφή)

.....

Αλεξία Πολιτάκη

Διπλωματούχος Μηχανικός Παραγωγής και Διοίκησης, Πολυτεχνείο Κρήτης

Διπλωματούχος Διατμηματικού Προγράμματος Μεταπτυχιακών Σπουδών Τεχνο-
οικονομικών Συστημάτων

Copyright © Αλεξία Πολιτάκη, 2021.

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.

Ευχαριστίες

Η παρούσα διπλωματική εργασία εκπονήθηκε κατά το ακαδημαϊκό έτος 2020-2021 στα πλαίσια της ολοκλήρωσης του Διατμηματικού Προγράμματος Μεταπτυχιακών Σπουδών (ΔΠΜΣ) "Τεχνο-Οικονομικά Συστήματα" που έχει οργανωθεί και υλοποιηθεί μέσα από τη συνεργασία της Σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών (ΗΜΜΥ) του Εθνικού Μετσόβιου Πολυτεχνείου (ΕΜΠ) και του Τμήματος Βιομηχανικής Διοίκησης και Τεχνολογίας του Πανεπιστημίου Πειραιώς.

Πριν την παρουσίαση και ανάλυση των αποτελεσμάτων της παρούσας διπλωματικής εργασίας, αισθάνομαι την ανάγκη να ευχαριστήσω ορισμένους ανθρώπους, χωρίς τη βοήθεια των οποίων δεν θα ήταν εφικτή η περάτωση της εν λόγω εργασίας.

Προσωπικά, θα ήθελα να εκφράσω τις ευχαριστίες μου στον επιβλέποντα Καθηγητή μου, κ. Βασίλειο Ασημακόπουλο, ο οποίος μου έδωσε την ευκαιρία να ασχοληθώ με τον πραγματικά ενδιαφέροντα τομέα των προβλέψεων στα πλαίσια της ερευνητικής δραστηριότητας της Μονάδας Προβλέψεων και Στρατηγικής.

Επίσης, θα ήθελα να ευχαριστήσω τον Καθηγητή κ. Ι. Ψαρρά και τον αναπληρωτή Καθηγητή, κ. Χ. Δούκα, για τη συμμετοχή τους στην επιτροπή εξέτασης της διπλωματικής εργασίας.

Θα ήθελα να ευχαριστήσω θερμά τον κ. Ευάγγελο Σπηλιώτη, ο οποίος ήταν παρών καθ' όλη τη διάρκεια εκπόνησης της διπλωματικής μου εργασίας, με πολύτιμες συμβουλές και παρατηρήσεις που έπαιξαν καθοριστικό ρόλο στην ολοκλήρωση της παρούσας εργασίας.

Κλείνοντας, θα ήθελα να ευχαριστήσω την οικογένεια μου για την απέραντη αγάπη και υποστήριξη που μου προσφέρουν, τα οποία μου δίνουν ώθηση για να πετυχαίνω τους στόχους μου.

Περίληψη

Σε ένα περιβάλλον πλήρως ανταγωνιστικό και αβέβαιο όπου οι συνθήκες συνεχώς μεταβάλλονται, οι επιχειρήσεις καλούνται να προσαρμοστούν αν θέλουν να επιβιώσουν και να αναπτυχθούν.

Στη βιομηχανία, μια παράμετρος που είναι κομβικής σημασίας για την ευημερία της επιχείρησης είναι η πρόβλεψη ζήτησης που καθορίζει τη λήψη αποφάσεων για τον προγραμματισμό της παραγωγής. Ειδικά στη βιομηχανία τροφίμων όπου τίθενται και παράγοντες αλλοιώσεων και περιορισμών αποθήκευσης απαιτείται πρόβλεψη ακόμα υψηλότερης ακρίβειας.

Στόχος της παρούσας διπλωματικής είναι η βελτιστοποίηση της πρόβλεψης ζήτησης προκειμένου να επιτευχθεί πιο στοχευμένος προγραμματισμός παραγωγής για την κάλυψη των αναγκών μιας μεσαίου μεγέθους εταιρείας τροφίμων που παράγει συσκευασμένα καταναλωτικά προϊόντα (CPG).

Σε πρώτο επίπεδο, η εργασία αναλύει σε θεωρητικό υπόβαθρο κάποιες έννοιες συνυφασμένες με την πρόβλεψη ζήτησης και τον προγραμματισμό παραγωγής ενώ στη συνέχεια παρατίθεται η μελέτη περίπτωσης όπου υλοποιούνται τα μοντέλα πρόβλεψης με πραγματικά δεδομένα με σκοπό την ανάδειξη της αποτελεσματικότητάς τους.

Συγκεκριμένα, μελετώνται ενδεικτικά κάποια στατιστικά μοντέλα πρόβλεψης ζήτησης και κάποια μοντέλα μηχανικής μάθησης που δεν παρέχονται από το εργαλείο πρόβλεψης ζήτησης που χρησιμοποιεί η εταιρεία. Βάσει των αποτελεσμάτων, αξιολογούνται τα οφέλη διασύνδεσης αυτών των μοντέλων με το ERP της επιχείρησης, με στόχο την αυτοματοποιημένη επιλογή της καλύτερης μεθόδου πρόβλεψης ανά brand προϊόντος για την πιο ακριβή παραγωγή προβλέψεων.

Τέλος, παρουσιάζονται τα αποτελέσματα των μεθόδων, παρατίθενται τα κύρια συμπεράσματα που προέκυψαν και προτείνονται πιθανά θέματα για μελέτη και ανάλυση στο μέλλον.

Λέξεις Κλειδιά: Χρονοσειρές, Μοντέλα προβλέψεων, Προγραμματισμός Παραγωγής, Τεχνικές Μηχανικής Μάθησης, Νευρωνικά Δίκτυα

Abstract

In a highly competitive and uncertain environment where conditions are constantly changing, companies are forced to adapt if they want to survive and thrive.

In any manufacturing business, one parameter that is crucial for the prosperity of the company is the demand forecasting that determines the decision making for the production planning. Especially in the food industry where the short perishability factors are a limitation to the maintenance of stocks, a forecast with a high accuracy level is required.

The scope of this thesis is to optimize demand forecasting in order to achieve more targeted production planning to meet the needs of a medium-sized food company that produces packaged consumer goods (CPG).

Initially, some concepts related to demand forecasting and production planning are analysed in a theoretical background, followed by the case study where the forecasting models are implemented with real data in order to examine their effectiveness.

Specifically, some statistical demand forecasting models and some machine learning models are examined that are not provided by the demand forecasting tool used by the company. Based on the results, it is under examination the benefits of connecting these models with the company's ERP. As a result, the goal of automating the selection of the best forecast method per brand resulting in more accurate forecast will be achieved.

Finally, the results of the methods are presented, the main conclusions that emerge are analysed and possible topics for future study are suggested.

Key words: Timeseries, Forecasting Techniques, Production Planning, Machine Learning Techniques, Neural Networks

Πίνακας Περιεχομένων

	Πίνακας Περιεχομένων	1
	Ευρετήριο Εικόνων	4
	Ευρετήριο Πινάκων	5
	Εισαγωγή	6
	Αντικείμενο της εργασίας	6
1	Δομή της εργασίας	7
1.1	Σχέση μεταξύ πρόβλεψης και προγραμματισμού παραγωγής	9
1.2		
1.3	Προβλέψεις	10
2	Εισαγωγή	10
2.1	Σημασία της πρόβλεψης ζήτησης στην εποχή μας	11
2.2		
2.3	Επιχειρηματική Ευφυΐα	15
2.4	Η μηχανική μάθηση στον προγραμματισμό παραγωγής	18
2.5	Κατηγορίες Μεθόδων Πρόβλεψης	20
2.6	Χρονοσειρά	22
2.6.1		
2.6.2	Εισαγωγή	22
2.6.3	Μοντέλο Χρονοσειρών	22
	Ανάλυση Χρονοσειράς	23
2.7	2.6.3.1 Γραφική Αναπαράσταση Χρονοσειράς	23
2.8	2.6.3.2 Ποιοτικά Χαρακτηριστικά της Χρονοσειράς	26
3	Βασικά στάδια στη διαδικασία πρόβλεψης	28
3.1	Παράγοντες που επηρεάζουν την επιλογή μεθόδου πρόβλεψης	30
3.1.1		
3.1.2	Μοντέλα πρόβλεψης	32
	Στατιστικά Μοντέλα	32
	Μέθοδος Naïve	33
	Seasonal Naïve	33

	Μοντέλα Παλινδρόμησης	33
	Μέθοδοι Εκθετικής Εξομάλυνσης	34
	3.1.4.1 Αλγόριθμος Εκθετικής Εξομάλυνσης ETS	34
3.1.3	Μέθοδος Theta	36
3.1.4	Μοντέλα ARIMA	38
	Συνδυασμός Μεθόδων	39
3.1.5		
3.1.6	Μοντέλα μηχανικής μάθησης	40
3.1.7	Νευρωνικά δίκτυα	42
3.2		
	3.2.1.1 Πολυεπίπεδο δίκτυο Perceptron (Multi-Layer Perceptron – MLP)	43
3.2.1	Δέντρα Απόφασης	45
	3.2.2 Στατιστικοί Δείκτες Σφάλματος	46
3.3	Συστήματα διαχείρισης επιχειρησιακών πόρων (Enterprise Resource Planning	
3.4	- ERP)	50
	3.4.1 Υποσύστημα προγραμματισμού παραγωγής	52
	3.4.2 Λύσεις λογισμικού ERP	53
	3.4.3 Προγραμματισμός Απαιτήσεων Υλικών (Materials Requirements	
3.4.4	Planning - MRP)	55
	3.4.5 Πως λειτουργεί το MRP	56
4	Ενσωματωμένα μοντέλα πρόβλεψης σε ERP συστήματα	59
4.1		
4.2	Πειραματική Διαδικασία	60
4.3	Καταγραφή του προβλήματος	60
4.4	Το R Studio ως εργαλείο προβλέψεων	61
4.4.1	Περιγραφή δεδομένων	63
4.4.2		
4.4.3	Ανάλυση Μεθοδολογίας	64
	Απαιτούμενες βιβλιοθήκες	64
	Διαχείριση κενών και αρνητικών τιμών	65
	Οπτικοποίηση δεδομένων	68

4.4.3.1	Διαγράμματα εποχιακότητας.....	72
	Επιλογή μεθόδων πρόβλεψης	73
	Αξιολόγηση Μεθόδων.....	74
	Αποτελέσματα στατιστικών μεθόδων.....	74
4.4.4	Αποτελέσματα μεθόδων μηχανικής μάθησης.....	79
4.5	Συμπεράσματα και Μελλοντικές Προεκτάσεις	83
4.5.1		
4.5.2	Συνοπτικά Συμπεράσματα.....	83
5	Προοπτικές και Μελλοντικές Προεκτάσεις	84
5.1	Βιβλιογραφία	85
5.2		

Ευρετήριο Εικόνων

Εικόνα 2.1 Ροή της αναλυτικής των δεδομένων.....	16
Εικόνα 2.2 Γνωστικά Αντικείμενα και εργαλεία ανάλυσης	17
Εικόνα 2.3 Οφέλη της χρήσης ΑΙ στις βιομηχανικές επιχειρήσεις-	18
Εικόνα 2.4 Κατηγορίες μεθόδων πρόβλεψης	20
Εικόνα 2.5 Κατηγοριοποίηση των ποσοτικών μεθόδων πρόβλεψης.....	21
Εικόνα 2.6 Μηνιαίες πωλήσεις σπιτιών για μονογονεϊκές οικογένειες στις ΗΠΑ.....	24
Εικόνα 2.7 Αριθμός επιβατών αυστραλιανής αεροπορικής εταιρείας.....	25
Εικόνα 2.8 Ημερήσια ζήτηση ηλεκτρικής ενέργειας.....	26
Εικόνα 2.9 Ποιοτικά χαρακτηριστικά της χρονοσειράς	26
Εικόνα 2.10 Τα βήματα της διαδικασίας πρόβλεψης	29
Εικόνα 3.1 Ροή διαδικασίας της Στατιστικής Πρόβλεψης	32
Εικόνα 3.2 Ταξινόμηση των μοντέλων εκθετικής εξομάλυνσης.....	35
Εικόνα 3.3 Αρχιτεκτονική Νευρωνικών Δικτύων	43
Εικόνα 3.4 Σύστημα Διαχείρισης Επιχειρησιακών Πόρων (ERP)	50
Εικόνα 3.5 Αντίκτυπο της πρόβλεψης στις λειτουργικές περιοχές της επιχείρησης..	51
Εικόνα 3.6 Τρόπος λειτουργίας του MRP	57
Εικόνα 3.7 Διαδικασία προγραμματισμού της παραγωγής	58
Εικόνα 4.1 Βασικά στάδια της πειραματικής διαδικασίας	64
Εικόνα 4.2 Εισαγωγή δεδομένων στην R	66
Εικόνα 4.3 Διαχωρισμός του συνόλου δεδομένων	67
Εικόνα 4.4 Απεικόνιση πολλαπλών χρονοσειρών	69
Εικόνα 4.5 Χρονοσειρά με εποχιακότητα.....	70
Εικόνα 4.6 Χρονοσειρά με εποχιακότητα και τάση	70
Εικόνα 4.7 Χρονοσειρά με πτωτική τάση.....	71
Εικόνα 4.8 Χρονοσειρά με πτωτική τάση και τυχειότητα	71
Εικόνα 4.9 Απεικόνιση χρονοσειράς σε διάγραμμα εποχιακότητας	72

Ευρετήριο Πινάκων

Πίνακας 4.1 Σφάλματα στατιστικών μεθόδων	75
Πίνακας 4.2 Σύγκριση σφαλμάτων στατιστικών μεθόδων με μέθοδο αυτόματης επιλογής	76
Πίνακας 4.3 Σύγκριση συνδυασμού μεθόδων(Combo) με μέθοδο αυτόματης επιλογής	79
Πίνακας 4.4 Συγκεντρωτικά σφάλματα μοντέλων MLP	80
Πίνακας 4.5 Σφάλματα μοντέλων MLP	81
Πίνακας 4.6 Σύγκριση σφαλμάτων όλων των μεθόδων	82

Εισαγωγή

Αντικείμενο της εργασίας

Όσο οι νέες τεχνολογίες αναπτύσσονται με ραγδαίους ρυθμούς τόσο πιο μεγάλη
1 απήχηση εμφανίζει και ο τομέας των τεχνικών προβλέψεων στις επιχειρήσεις. Οι
1.1 προβλέψεις που προέκυπτan αποκλειστικά από την εμπειρία των στελεχών μιας
εταιρείας έχουν περιοριστεί και πλέον λειτουργούν επικουρικά σε συνδυασμό με
στατιστικές μεθόδους που αναλύουν και επεξεργάζονται τα δεδομένα. Η αξιοποίηση
σύγχρονων εργαλείων ανάλυσης δεδομένων κερδίζει συνεχώς έδαφος στις
επιχειρήσεις. Όμως, παρόλο που ζούμε στην εποχή της 4ης Βιομηχανικής
Επανάστασης, η ανθρώπινη «διαίσθηση» και η «εμπειρία» παραμένουν κυρίαρχα
κριτήρια των επιχειρηματικών επιλογών στην Ελλάδα, καθώς τα ψηφιακά εργαλεία δεν
συνοδεύονται από τον ανάλογο διοικητικό και λειτουργικό μετασχηματισμό. Σύμφωνα
με πρόσφατη έρευνα της Deloitte, μόνο το 39% των επιχειρήσεων διαθέτουν
διοικητικές διαδικασίες βασισμένες σε κουλτούρα και συστήματα ανάλυσης
δεδομένων ενώ τα 2/3 των επιχειρήσεων περιορίζουν την συλλογή και ανάλυση
δεδομένων σε μια πολύ μικρή ομάδα εργαζομένων. Τα οφέλη είναι ήδη ορατά σε
εκείνες τις επιχειρήσεις που επενδύουν σε τεχνολογίες ανάλυσης δεδομένων καθώς η
βιομηχανία έχει τη δυνατότητα να προσαρμόσει την ποσότητα παραγωγής βάσει
προβλεπόμενης ζήτησης, να διαχειριστεί το ύψος των αποθεμάτων, να διενεργήσει
προληπτική συντήρηση, να βελτιστοποιήσει το μίγμα πωλήσεων ανά κανάλι
προώθησης, κ.λπ. Στην παρούσα διπλωματική εξετάζεται η άμεση σχέση που έχει η
πρόβλεψη ζήτησης με τον προγραμματισμό παραγωγής.

Αντικείμενο, λοιπόν, της παρούσας διπλωματικής είναι η βελτιστοποίηση της
πρόβλεψης ζήτησης προκειμένου οι υπεύθυνοι παραγωγής να είναι σε θέση να
παίρνουν πιο στοχευμένες αποφάσεις που αφορούν στον προγραμματισμός παραγωγής
για την κάλυψη των αναγκών μιας μεσαίου μεγέθους εταιρείας τροφίμων που παράγει
συσκευασμένα καταναλωτικά προϊόντα (CPG). Συγκεκριμένα, θα μελετηθούν τα
οφέλη κάποιων στατιστικών μοντέλων πρόβλεψης και κάποιων μοντέλων μηχανικής
μάθησης που δεν παρέχονται από το εργαλείο πρόβλεψης ζήτησης που χρησιμοποιεί η
εταιρεία. Ανάλογα με τα οφέλη θα αξιολογηθεί το σενάριο σύνδεσης αυτών των
μοντέλων με το ERP της επιχείρησης, με στόχο την αυτοματοποιημένη επιλογή της
καλύτερης μεθόδου πρόβλεψης ανά brand προϊόντος για πιο ακριβή παραγωγή
προβλέψεων.

Δομή της εργασίας

Η παρούσα διπλωματική εργασία αποτελείται από πέντε κεφάλαια.

- 1.2 Στο πρώτο κεφάλαιο γίνεται μια εισαγωγή του θέματος της παρούσας διπλωματικής εργασίας και της βασικής δομής των κεφαλαίων της.

Στο δεύτερο κεφάλαιο περιγράφεται η πρόβλεψη, η εξάρτηση μεταξύ πρόβλεψης ζήτησης και προγραμματισμού παραγωγής και η σημασία της πρόβλεψης στις σύγχρονες βιομηχανικές μονάδες. Επίσης, γίνεται μια εισαγωγή στον όρο της επιχειρηματικής ευφυΐας, τη σχέση του με την μηχανική μάθηση και πώς μπορούν αυτές οι έννοιες που παρέχουν συνεχώς νέα εργαλεία να αξιοποιηθούν στον χώρο των επιχειρήσεων και συγκεκριμένα στον προγραμματισμό παραγωγής. Σε αυτό το κεφάλαιο περιγράφονται, επίσης, οι κατηγορίες μεθόδων πρόβλεψης, τα μοντέλα χρονοσειρών, καθώς και τα βασικά χαρακτηριστικά τους.

Στο τρίτο κεφάλαιο αναλύονται παρουσιάζονται κάποια κύρια στατιστικά μοντέλα και μοντέλα μηχανικής μάθησης και αναλύονται σε βάθος οι μέθοδοι πρόβλεψης που χρησιμοποιήθηκαν στην παρούσα διπλωματική. Σε αυτό το κεφάλαιο παρουσιάζονται επίσης τα στατιστικά σφάλματα που χρησιμοποιούνται για την αξιολόγηση των μεθόδων πρόβλεψης. Το τέλος του κεφαλαίου αφιερώνεται στα συστήματα επιχειρησιακών πόρων(ERP) καθώς και του κομβικού ρόλου που διαδραματίζουν τα ενσωματωμένα συστήματα MRP στην αποτελεσματικότητα της παραγωγικής λειτουργίας. Αναλύονται τα μοντέλα πρόβλεψης που περιλαμβάνει το κάθε εργαλείο ERP προκειμένου να προβλέψει τη ζήτηση που θα χρησιμοποιηθεί σαν είσοδος για την εξαγωγή του βασικού χρονοδιαγράμματος παραγωγής.

Το τέταρτο κεφάλαιο αποτελείται από το case-study που αφορά πραγματικά δεδομένα πωλήσεων μιας ελληνικής εταιρείας που δραστηριοποιείται στο χώρο των σνακ κατέχοντας ηγετική θέση. Οι πωλήσεις αφορούν τα επιμέρους brands που παράγει η εταιρεία. Παρουσιάζεται η μεθοδολογία που ακολουθήθηκε στη διπλωματική και γίνεται αναφορά στην επεξεργασία των δεδομένων, την υλοποίηση των μοντέλων πρόβλεψης και τη διαδικασία σύγκρισης των μεθόδων. Αναλύονται στατιστικά μοντέλα αλλά και μοντέλα μηχανικής μάθησης.

Στόχος είναι η ανάδειξη της αποτελεσματικότητας των μεθόδων που εξετάστηκαν, μέσω της αξιολόγησης τους από πραγματικά δεδομένα.

Στο πέμπτο και τελευταίο κεφάλαιο παρατίθενται τα κύρια συμπεράσματα που προέκυψαν με βάση τα αποτελέσματα των πειραμάτων και προτείνονται πιθανά θέματα για μελέτη και ανάλυση στο μέλλον.

Σχέση μεταξύ πρόβλεψης και προγραμματισμού παραγωγής

1.3 Ενώ και οι δυο έννοιες σχετίζονται με μελλοντικά γεγονότα, η πρόβλεψη προηγείται του προγραμματισμού, αποτελώντας τη βάση του και διαδραματίζοντας καθοριστικό ρόλο στη διαδικασία σχεδιασμού του. Υπάρχει μια ξεκάθαρη σύνδεση μεταξύ τους και έχει αποδειχθεί ότι μια ακριβής πρόβλεψη μπορεί να αποδυναμωθεί αν χρησιμοποιηθεί σε ένα πλαίσιο φτωχού και ελλιπούς προγραμματισμού.

Ο προγραμματισμός, περιλαμβάνει τον καθορισμό των κατάλληλων ενεργειών που απαιτούνται για την μελλοντική πορεία δράσης, προκειμένου να επιτευχθούν οι επιθυμητοί στόχοι του οργανισμού. Συνεπώς, μέσω του προγραμματισμού, γίνεται μια συνειδητή προσπάθεια να τροποποιηθούν μελλοντικά γεγονότα. Η πρόβλεψη, λαμβάνοντας υπόψη τις προηγούμενες και τις τρέχουσες επιδόσεις και δεδομένα, εκτιμάει τη απόδοση του οργανισμού. Ο προγραμματισμός τονίζει τα γεγονότα και τις προσδοκίες. Η πρόβλεψη, από την άλλη πλευρά, βασίζεται σε γεγονότα και ορισμένες παραδοχές με βάση την προηγούμενη και την τρέχουσα απόδοση.

Πολλές φορές, υπάρχει σύγχυση ανάμεσα στους δυο όρους, δίνοντας αυξημένη βαρύτητα στον προγραμματισμό, ωστόσο οι μελέτες αλλά και πραγματικά δεδομένα επιχειρήσεων, υποδεικνύουν ότι και οι δυο έννοιες είναι ζωτικής σημασίας για την ανάπτυξη μιας εταιρείας με τον προγραμματισμό να αποτελεί την αντίδραση στις προβλέψεις έτσι ώστε να επιτευχθούν οι στόχοι που έχουν τεθεί.

Προβλέψεις

Εισαγωγή

Ως πρόβλεψη, ορίζεται η διαδικασία συγκέντρωσης όλης της διαθέσιμης πληροφορίας, τόσο των ιστορικών δεδομένων, όσο και γνωστών μελλοντικών δράσεων που σχετίζονται με το αντικείμενο προς μελέτη και η αξιοποίησή της για την παραγωγή ισχυρισμών για γεγονότα, των οποίων τα πραγματικά αποτελέσματα συνήθως δεν έχουν ακόμα παρατηρηθεί. Οι συνεχώς μεταβαλλόμενες επιχειρηματικές συνθήκες ως αποτέλεσμα του παγκόσμιου ανταγωνισμού και των ραγδαίων τεχνολογικών αλλαγών ωθούν τις επιχειρήσεις για όσο το δυνατόν πιο ακριβείς προβλέψεις. Το ενδιαφέρον αυτό, προκύπτει από την ανάγκη λήψης απόφασης αντιμετωπίζοντας την αβεβαιότητα του μέλλοντος. Βέβαια είναι κομμάτι της ανθρώπινης φύσης να επιζητούμε το σίγουρο, και το γεγονός ότι η πρόβλεψη είναι συνυφασμένη με τις πιθανότητες προκαλεί δισταγμό και καχυποψία. Ωστόσο, όταν το ευρύ κοινό συνειδητοποιήσει ότι οι προβλέψεις είναι κομμάτι της ζωής μας θα καταφέρει να κατανοήσει καλύτερα, να εξοικειωθεί και να εκτιμήσει αυτόν τον τόσο σημαντικό τομέα. Ένα ενδεικτικό παράδειγμα της επικαιρότητας είναι, η κατάσταση που επικρατεί σχετικά με την πανδημία του ιού COVID-19, η οποία έχει κάνει ακόμα και τους πιο δύσπιστους να αντιληφθούν ότι ζούμε σε μια εποχή με τεράστιες αβεβαιότητες και μέσω των προβλέψεων θα καταφέρουμε να τις αντιμετωπίσουμε.

Από τεχνολογικής άποψης, οι επιχειρήσεις έχουν όλα τα εφόδια για να προβούν σε όσο το δυνατόν πιο έγκυρες προβλέψεις καθώς η συλλογή δεδομένων αλλά κυρίως η αποθήκευση και επεξεργασία δεδομένων έχουν σημειώσει ραγδαία βελτίωση την τελευταία δεκαετία. Αξιοποιώντας λοιπόν, αυτές τις τρεις παραμέτρους, η πρόβλεψη ζήτησης είναι έτοιμη να προαχθεί σε ένα κομβικό ρόλο στην αλυσίδα αξίας μιας επιχείρησης. Οι προβλέψεις αποτελούν αναπόσπαστο κομμάτι μιας επιχείρησης προκειμένου να καθορίσει τις προβλεπόμενες αναγκαίες πηγές, να προγραμματίσει τις υπάρχουσες πηγές και να αποκτήσει τις υπόλοιπες που θα χρειαστούν για την παραγωγή. Οι ακριβείς προβλέψεις επιτρέπουν στις επιχειρήσεις να παίρνουν εύστοχες στρατηγικές αποφάσεις, να χρησιμοποιούν αποτελεσματικά τις δυνατότητες των μηχανημάτων, να μειώνουν τους χρόνους παραγωγής και τα αποθέματα παρέχοντας τους ανταγωνιστικό πλεονέκτημα.

Παρότι η τεχνολογία έχει εξελιχθεί αρκετά, ακόμα υπάρχουν αρκετές επιχειρήσεις που δεν έχουν πειστεί για την αξία των προβλέψεων. Μια πρόκληση που έχει να αντιμετωπίσει η επιστήμη των προβλέψεων είναι, ότι η εποχή μας παράγει μεγάλο όγκο δεδομένων δυσκολεύοντας πολλές φορές την εξαγωγή χρήσιμων πληροφοριών από αυτά. Επίσης, αρκετές επιχειρήσεις αντιμετωπίζουν με σκεπτικισμό τον τομέα των προβλέψεων λόγω της αδυναμίας των μεθόδων να προμηνύσουν έγκαιρα επερχόμενες αλλαγές. Στην πραγματικότητα όμως, τότε περισσότερο από πότε υπάρχει η ουσιαστική ανάγκη για προβλέψεις και αυτό επιβεβαιώνεται από περιόδους γενικευμένων κρίσεων που έχουν αυξήσει κατακόρυφα τη ζήτηση για πραγματοποίηση προβλέψεων. Έχει παρατηρηθεί, ειδικά στην Ελλάδα, ότι ένα μεγάλο ποσοστό επιχειρήσεων δεν δίνουν την απαιτούμενη βαρύτητα στον τομέα των προβλέψεων. Εν τούτοις, δεν παύει να είναι οξύμωρο το γεγονός ότι οι επιχειρήσεις, χρησιμοποιούν γνώσεις και εργαλεία από εξειδικευμένους επιστημονικούς τομείς για να κατασκευάσουν ένα προϊόν, σύμφωνα με δομημένες οδηγίες και προδιαγραφές, εν αντιθέσει με την πρόβλεψη ζήτησης του ίδιου προϊόντος για την οποία επιστρατεύεται κατά κύριο λόγο το ένστικτο. Παρ' όλα αυτά, κανείς δε θα μπορούσε να αμφισβητήσει την αξία της κριτικής και οξείας σκέψης του ανθρώπινου παράγοντα ως αποτέλεσμα πολυετούς ενασχόλησης με τον εκάστοτε τομέα. Αυτός είναι και από τους σημαντικότερους λόγους ύπαρξης συνεχούς βελτίωσης αλλά και δημιουργίας των μεθόδων πρόβλεψης.

Εν κατακλείδι, η αναγκαιότητα της πρόβλεψης είναι αναμφισβήτητη και οι προβλέψεις θα έπρεπε να αποτελούν καθοριστικό κομμάτι οποιασδήποτε επιχείρησης διότι η ακριβέστερη πρόβλεψη μελλοντικών καταστάσεων θα οδηγήσει σε ορθές στρατηγικές

2.2 αποφάσεις, οι οποίες συνεπάγονται και βέλτιστο κέρδος για τις επιχειρήσεις.

Σημασία της πρόβλεψης ζήτησης στην εποχή μας

Είναι γεγονός ότι το επιχειρηματικό τοπίο έχει αλλάξει δραματικά τα τελευταία χρόνια. Αυτές οι αλλαγές που, μεταξύ άλλων, σχετίζονται με την οικονομική ύφεση και τον αυξημένο ανταγωνισμό, έχουν φέρει τις επιχειρήσεις αντιμέτωπες με ραγδαίες αυξήσεις σε πρώτες ύλες, ενέργεια και κόστη μεταφοράς. Με την αύξηση του ανταγωνισμού, γίνεται ακόμα πιο επιτακτική η ανάγκη για τη δημιουργία ισορροπίας μεταξύ προσφοράς και ζήτησης.

Η πρόβλεψη ζήτησης αποτελεί για κάθε επιχείρηση τη βάση του στρατηγικού σχεδιασμού της παραγωγής, των πωλήσεων και των χρηματοοικονομικών της. Χρησιμοποιείται στον προγραμματισμό και τον έλεγχο της παραγωγής τόσο βραχυπρόθεσμα όσο και μακροπρόθεσμα. Αν υπάρχει κάτι που οι εργαζόμενοι που στελεχώνουν τις θέσεις διοίκησης θα επιθυμούσαν, είναι μια κρυστάλλινη σφαίρα. Μια μαγική ικανότητα δηλαδή, να βολιδοσκοπούν το μέλλον προκειμένου να μειώσουν την πολυπλοκότητα και την αβεβαιότητα στις σύγχρονες επιχειρήσεις και να παρέχουν μια διαδρομή σταθερότητας και βεβαιότητας σε μια πολύ-παραγοντική αλυσίδα αξίας. Ενώ μια κρυστάλλινη σφαίρα δεν αποτελεί επιλογή, η διοίκηση κατέχει ένα καίριο εργαλείο ως σύμμαχο στον προσδιορισμό μελλοντικών αναγκών παραγωγής, στη βέλτιστη διαχείριση αποθέματος καθώς και στις στρατηγικές κατανομής θέσεων εργασίας με στόχο τη μέγιστη αποδοτικότητα και παραγωγικότητα, τις προβλέψεις. Μέσω των προβλέψεων, επιτρέπεται μια πιο στοχευμένη στρατηγική και τακτικός σχεδιασμός για την κατανομή των πόρων στα επιμέρους προϊόντα της εταιρείας οδηγώντας σε αυξημένη κερδοφορία. Στην πραγματικότητα, η πρόβλεψη είναι μια ολοένα και πιο πολύτιμη πρόταση για τις βιομηχανίες, που μια μελέτη του 2016 από την Gartner έδειξε ότι η πρόβλεψη (και η ακρίβεια αυτών) και η μεταβλητότητα της ζήτησης ήταν δύο από τα μεγαλύτερα εμπόδια που αντιμετωπίζουν οι βιομηχανίες κατά την εποπτεία των ροών της εφοδιαστικής τους αλυσίδας.

Στην τρέχουσα παράγραφο θα δοθούν εν συντομία τα πλεονεκτήματα που παρέχει η πρόβλεψη ζήτησης για τον προγραμματισμό και την παραγωγή στο σύγχρονο βιομηχανικό τοπίο.

Αποτελεσματικότερος προγραμματισμός παραγωγής

Μέσω των προβλέψεων, οι επιχειρήσεις βλέποντας στο μέλλον μπορούν να επιτύχουν την ικανοποίηση των απαιτήσεων των πελατών τους και της αγοράς και να ευθυγραμμιστούν με τη διαθεσιμότητα πρώτων υλών και εξαρτημάτων. Με αυτό τον τρόπο, μπορούν να λειτουργούν με περισσότερη ευελιξία, διαφάνεια και να προσαρμόζονται άμεσα σε μεταβαλλόμενα περιβάλλοντα παραγωγής.

Βέλτιστη διαχείριση και μείωση αποθεμάτων

Όταν ο κατασκευαστής κατανοήσει καλύτερα και προβλέψει τη ζήτηση, μόνο τότε μπορεί να συνεργαστεί αποτελεσματικότερα με τους προμηθευτές για να επιτύχει τα

βέλτιστα επίπεδα αποθέματος μειώνοντας τη πιθανότητα για ελλείψεις ή πλεονάσματα. Οι δυνατότητες που παρέχει η πρόβλεψη δίνουν στον κατασκευαστή μια πλήρη εικόνα της εφοδιαστικής αλυσίδας προκειμένου να αξιολογήσει με ακρίβεια το επίπεδο της ζήτησης των πελατών σε σχέση με τον όγκο του αποθέματος που είναι απαραίτητα για την ικανοποίηση των παραγγελιών και τη διασφάλιση των προγραμματισμένων παραδόσεων. Η μείωση του αποθέματος, όχι μόνο περιορίζει τις περιττές διαδικασίες και εξοικονομεί χώρο αλλά βοηθάει τις εταιρείες να εναρμονίσουν τις υφιστάμενες διαδικασίες τους εξαλείφοντας δαπανηρές ζημίες που είναι σύμφυτες με ένα αχρησιμοποίητο απόθεμα στην αποθήκη. Σύμφωνα με το ερευνητικό και συμβουλευτικό ινστιτούτο Gartner, εταιρίες που εφαρμόζουν εξειδικευμένες πρακτικές πρόβλεψης έχουν τη δυνατότητα να μειώσουν το απόθεμά τους έως και 15%.

Μείωση κόστους

Παραπάνω, αναλύθηκε πώς η πρόβλεψη μειώνει το κόστος που σχετίζεται με αχρησιμοποίητα υλικά ή εξαρτήματα, αλλά η πρόβλεψη συμβάλλει επίσης στη μείωση του κόστους που σχετίζεται με μια σειρά άλλων κρίσιμων εργασιών της παραγωγής όπως η διαχείριση των ανθρώπινων πόρων για την ανάθεση εργασιών. Καθώς η πρόβλεψη επηρεάζει όλο τον κύκλο της παραγωγικής διαδικασίας, από την αρχή έως το τέλος και κατά συνέπεια όλη την αλυσίδα αξίας της βιομηχανικής μονάδας, μια αποτελεσματική και οικονομικά μονάδα παραγωγής σημαίνει μια πιο ισχυρή βιομηχανία. Σύμφωνα με τον ΣΕΒ, η αύξηση των εσόδων μπορεί να ανέλθει μέχρι 3%.

Αποτελεσματικότερη εφοδιαστική των μεταφορών (transport logistics)

Σε πολλές επιχειρήσεις, υπάρχουν αρκετά κρυμμένα κόστη στον ευρύτερο τομέα της εφοδιαστικής αλυσίδας και συγκεκριμένα στις εμπορευματικές μεταφορές. Προκειμένου να ομαλοποιηθεί ή ακόμα και να περιοριστεί αυτό το κόστος, κάθε εταιρεία προσπαθεί να συνδυάσει αποστολές ή να δοκιμάσει νέες μεθόδους διαμετακόμισης, σε μερικές περιπτώσεις ακόμα και να αλλάξει τις ημερομηνίες παράδοσης σύμφωνα με τη ζήτηση του πελάτη. Ενώ οι παραπάνω αποτελούν πιθανές επιλογές, οι προβλέψεις επιτρέπουν στις εταιρείες να κινηθούν ένα βήμα παραπέρα από τις συμβατικές επιλογές έτσι ώστε να αναλύσουν σε βάθος τη στρατηγική των μεταφορών τους εντοπίζοντας συγκεκριμένα πεδία που έχουν περιθώρια βελτίωσης. Επειδή η αποτελεσματικότερη εφοδιαστική των μεταφορών αφορά στον ταχύτερο και

πιο αποτελεσματικό τρόπο μεταφοράς προϊόντων από το σημείο Α στο σημείο Β, η πρόβλεψη δίνει στις εταιρείες τη δυνατότητα να δουν πότε, πού, πώς και γιατί πρέπει να εκτελεστούν κάποιες αποφάσεις στρατηγικής σημασίας που αφορούν στις μεταφορές και την αξία που προσδίδουν αυτές οι ενέργειες στην εφοδιαστική αλυσίδα.

Αποτελεσματικότερη επικοινωνία εντός της εταιρείας

Καθώς όλοι οι εσωτερικοί ενδιαφερόμενοι αρχίζουν να εμπιστεύονται τη διαδικασία πρόβλεψης, ευθυγραμμίζεται η ροή των εργασιών, οδηγώντας σε μια ποιοτικότερη συνεργασία μεταξύ των επιμέρους τμημάτων. Η οικοδόμηση ποιοτικών σχέσεων, μεταφράζεται σε ενισχυμένη ενοποίηση του δικτύου που συμβάλλει στην ελαχιστοποίηση των πιέσεων της δυναμικής της αγοράς που περιβάλλει τα προϊόντα της επιχείρησης. Λόγω της βελτιωμένης συνεργασίας, τα ανώτερα στελέχη αντιλαμβάνονται καλύτερα τους παράγοντες που ωθούν την κερδοφορία, με αποτέλεσμα μεγαλύτερο έλεγχο του προϋπολογισμού και πιο στοχευμένες επενδύσεις.

Οι τεχνικές προβλέψεις απαντώνται σε μια πληθώρα εφαρμογών στον ευρύτερο τομέα της παραγωγής. Ενδεικτικά, παρουσιάζονται παρακάτω κάποιες από αυτές:

- Πωλήσεις - Εισαγωγή νέου προϊόντος
- Μάρκετινγκ – Προωθητικές ενέργειες
- Προληπτική Συντήρηση
- Προγραμματισμός Εργατικού Δυναμικού
- Προμήθειες και Αξιοποίηση Παγίων

Επιχειρηματική Ευφυΐα

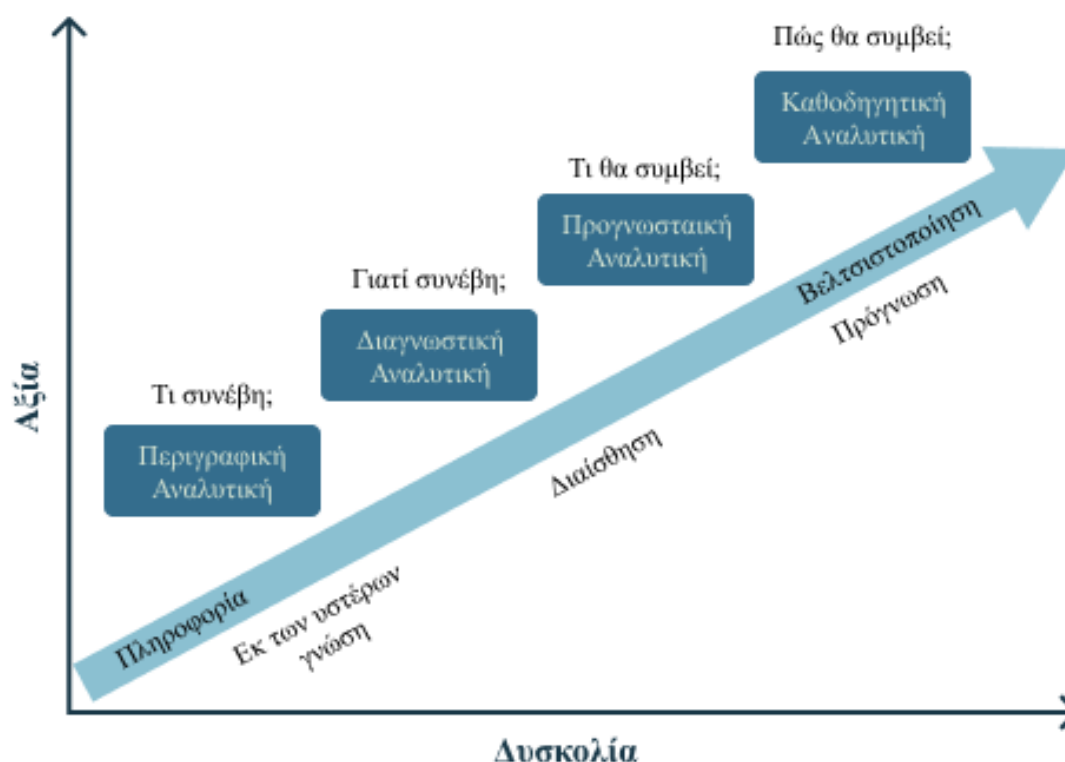
2.3 Η ανάγκη των επιχειρήσεων για βελτιωμένη πληροφόρηση και για πιο στοχευμένη και οργανωμένη διαδικασία λήψης αποφάσεων, η διαθεσιμότητα μεγάλου όγκου και ποικίλου περιεχομένου δεδομένων, και η ανάπτυξη νέων τεχνολογιών και μεθοδολογιών για την ανάλυση των δεδομένων, αποτέλεσαν τα εφαλτήρια για την ανάπτυξη των συστημάτων Επιχειρηματικής Ευφυΐας. Στη βιβλιογραφία απαντώνται διάφοροι ορισμοί της επιχειρηματικής ευφυΐας με τον πιο ολοκληρωμένο να την ορίζει ως ένα σύνολο από μεθόδους ανάλυσης, τεχνολογίες, ικανότητες και στρατηγικές, οι οποίες στόχο έχουν την επεξεργασία των διαθέσιμων δεδομένων και την εξαγωγή χρήσιμης πληροφορίας από αυτά, για την υποστήριξη της διαδικασίας λήψης επιχειρηματικών αποφάσεων. Η ανάπτυξη τέτοιων συστημάτων αποτελεί μια πολύ ισχυρή και σύγχρονη τάση εφαρμογής τεχνολογιών πληροφορικής στις σημερινές επιχειρήσεις. Μέσα στα εργαλεία που περιλαμβάνουν είναι και συστήματα λογισμικού για επιχειρηματικές αναλύσεις «Business Analytics». Τα Business Analytics διευρύνουν την επιχειρηματική ευφυΐα με την οπτική του μέλλοντος και βασίζονται κυρίως στη στατιστική ανάλυση επιχειρηματικών δεδομένων. Μπορούν να δώσουν απαντήσεις σε ερωτήματα σχετικά με τους λόγους, τις επιδράσεις, τις αλληλεπιδράσεις ή τις συνέπειες γεγονότων. Μπορούν ακόμα να εξετάσουν διάφορα σενάρια και να προτείνουν εναλλακτικούς τρόπους δράσης τα οποία χωρίζονται σε τέσσερα στάδια.

Περιγραφική Αναλυτική(Descriptive Analytics): Αποτελεί το πρώτο στάδιο της Επιχειρηματικής Αναλυτικής και αφορά στην ανάλυση των δεδομένων με στόχο την κατανόηση της παρούσας κατάστασης μιας επιχείρησης. Το αποτέλεσμα αυτής της φάσης είναι συνήθως κάποια αναφορά με πίνακες και διαγράμματα.

Διαγνωστική Αναλυτική (Diagnostic Analytics): Αυτός ο τύπος χρησιμοποιεί τεχνικές όπως ανίχνευση δεδομένων, εξόρυξη δεδομένων και συσχετισμούς για την αποκάλυψη των βασικών αιτιών των γεγονότων.

Προγνωστική Αναλυτική (Predictive Analytics): Οι προγνωστικές αναλύσεις προβλέπουν τη δυνατότητα μελλοντικών γεγονότων χρησιμοποιώντας στατιστικά μοντέλα και τεχνικές μηχανικής μάθησης(machine learning). Αυτός ο τύπος βασίζεται στα αποτελέσματα της περιγραφικής λυσης για την ανάπτυξη μοντέλων που μπορούν να εξάγουν την πιθανότητα εμφάνισης των επιθυμητών αποτελεσμάτων.

Καθοδηγητική Αναλυτική (Prescriptive Analytics): Η καθοδηγητική αναλυτική πηγαίνει ένα βήμα πιο πέρα από τη πρόγνωση, παρέχοντας συστάσεις για τις επόμενες βέλτιστες ενέργειες και επιτρέποντας την πιθανή διαχείριση των μελλοντικών γεγονότων με στόχο τη βελτιστοποίηση των αποτελεσμάτων για την επιχείρηση.



Εικόνα 2.1 Ροή της αναλυτικής των δεδομένων

Τα γνωστικά αντικείμενα που υποστηρίζουν διαδικασίες Business Analytics μέσω μεθόδων ανάλυσης δεδομένων όπως, ενδεικτικά, η τεχνητή νοημοσύνη, η βαθιά εκμάθηση, η μηχανική μάθηση επικαλύπτονται σε μεγάλο βαθμό. Ωστόσο, αρκετές είναι και οι μεταξύ τους διαφορές. Η βάση δεδομένων που διαθέτει μια επιχείρηση, αποτελεί τον ακρογωνιαίο λίθο για την ανάπτυξη εφαρμογών ανάλυσης δεδομένων και αλγορίθμων μηχανικής μάθησης. Σε επόμενο στάδιο, εφόσον η επιχείρηση έχει διαθέσιμα δεδομένα μπορεί να εκτελεί διαδικασίες με «ευφυή τρόπο». Στην αρχή, ενδείκνυται η χρήση κάποιων απλών στατιστικών μεθόδων οι οποίες είναι κατανοητές, μπορούν να δώσουν μια πρώτη εικόνα για το πώς συσχετίζονται οι μεταβλητές και συμβάλλουν στην αναγνώριση κάποιων μοτίβων. Έπειτα, μπορεί να ξεκινήσει η διαδικασία της μοντελοποίησης μέσω αλγορίθμων μηχανικής μάθησης και πιο

σύνθετων μεθόδων, με τελευταίο στάδιο εξέλιξης τις εφαρμογές τεχνητής νοημοσύνης. Η τεχνητή νοημοσύνη (AI) είναι ένας τομέας της επιστήμης των υπολογιστών που ασχολείται με τη σχεδίαση και την υλοποίηση προγραμμάτων τα οποία είναι ικανά να μαθαίνουν από την εμπειρία, να προσαρμόζονται σε νέα εισαγόμενα δεδομένα και να εκτελούν λειτουργίες που σχετίζονται με την ανθρώπινη λογική ή σκέψη, ακόμα και να μπορούν να διαμορφώσουν μια δική τους λογική.



Εικόνα 2.2 Γνωστικά Αντικείμενα και εργαλεία ανάλυσης

Σύμφωνα με τον ΣΕΒ, οι εφαρμογές τεχνητής νοημοσύνης στις βιομηχανικές επιχειρήσεις μπορούν να βελτιώσουν έως 5% την αποδοτικότητα στην παραγωγή, ως 13% την κερδοφορία, έως 12% την εξοικονόμηση καυσίμων, έως 10% το χρόνο διάθεσης προϊόντων στην αγορά. Στο λιανικό εμπόριο, επιτυγχάνεται έως 20% μείωση στα αποθέματα, έως 30% μείωση του χρόνου αποθεματοποίησης και έως 30% αύξηση των online πωλήσεων. Στην ηλεκτρική ενέργεια, η τεχνητή νοημοσύνη επιφέρει έως 20% αύξηση της παραγωγής και έως 20% βελτίωση της κερδοφορίας. Με τα κατάλληλα μοντέλα μηχανικής μάθησης, οι οργανισμοί έχουν τη δυνατότητα να προβλέπουν συνεχώς τις αλλαγές στο εσωτερικό τους, ώστε να είναι σε θέση να προβλέψουν τι θα ακολουθήσει.



Εικόνα 2.3 Οφέλη της χρήσης AI στις βιομηχανικές επιχειρήσεις-
Analytics & Τεχνητή Νοημοσύνη (sev.org.gr)

Η μηχανική μάθηση στον προγραμματισμό παραγωγής

2.4

Η μηχανική μάθηση είναι ένας κλάδος του ευρύτερου πεδίου της τεχνητής νοημοσύνης (AI) που ασχολείται με το σχεδιασμό και την υλοποίηση αλγορίθμων οι οποίοι μαθαίνουν από δεδομένα χωρίς να έχουν προγραμματιστεί συγκεκριμένα για αυτό που μαθαίνουν. Τα συστήματα μπορούν να μάθουν και να βελτιωθούν με την εμπειρία ώστε να είναι σε θέση να βελτιστοποιήσουν ένα μοντέλο το οποίο μπορεί να χρησιμοποιηθεί για την επίλυση μελλοντικών περιπτώσεων.

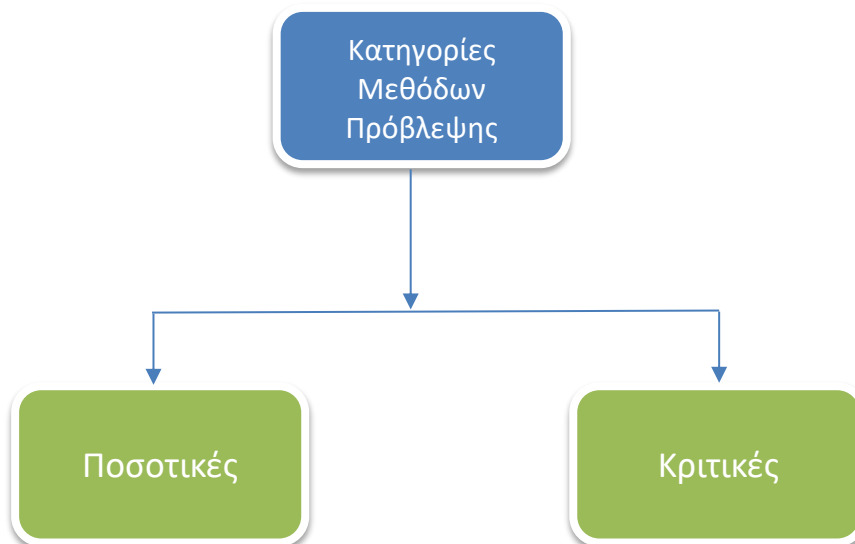
Ο προγραμματισμός παραγωγής είναι ένα κρίσιμο εργαλείο σε μια βιομηχανία και μπορεί να έχει άμεσο αντίκτυπο στην αύξηση ή μείωση της παραγωγικότητας. Υπό αυτό το πρίσμα, η χρήση της μηχανικής μάθησης στον προγραμματισμό παραγωγής στο πλαίσιο της 4^{ης} Βιομηχανικής Επανάστασης, μπορεί να αποδειχτεί πολύ καρποφόρα, ειδικά σε μια βιομηχανία τροφίμων όπου εμπλέκει πολύπλοκες διεργασίες και υπάρχει έντονη μεταβλητότητα προκαλώντας συχνά μια μεγάλη απόκλιση μεταξύ της πραγματικής και προγραμματισμένης παραγωγής. Οι εν λόγω αποκλίσεις έχουν οικονομικό κόστος για την εταιρεία. Η χρήση αλγορίθμων μηχανικής μάθησης αναδεικνύει καινοτόμες λύσεις αντιμετώπισης της αβεβαιότητας στον σχεδιασμό παραγωγής που βελτιώνουν τις κλασικές μεθόδους και παρέχουν στους διαχειριστές παραγωγής πολύτιμες πληροφορίες βελτιώνοντας δυνητικά την κερδοφορία της επιχείρησης. Σε αυτές τις δραστηριότητες περιλαμβάνεται η προληπτική συντήρηση, η

παρακολούθηση αποθεμάτων, η πρόβλεψη ζήτησης, κ.λπ. Άλλος ένας λόγος για τον οποίο η μηχανική μάθηση έχει βρει ανταπόκριση σε τόσα πεδία εφαρμογών είναι ότι πλέον οι επιχειρήσεις δεν είναι απαραίτητο να έχουν στο δυναμικό τους πολλά εξειδικευμένα στελέχη με πολυετή εμπειρία που να γνωρίζουν την αγορά και τη βιομηχανία, καθώς μέσα από τη χρήση των δεδομένων, ένας έμπειρος αναλυτής μπορεί να παρέχει στοχευμένες λύσεις. Φυσικά, όπως έχει προαναφερθεί, δεν υφίσταται η έννοια της μηχανικής μάθησης χωρίς την ύπαρξη δεδομένων, επομένως, η έκρηξη των δεδομένων και η εισαγωγή της χρήσης κινητών και μικροσυσκευών που συλλέγουν δεδομένα ευνοούν τη χρήση μηχανικής μάθησης για την επίλυση καθημερινών προβλημάτων στην παραγωγή. Άλλος ένας παράγοντας που καθιστά αυτές τις λύσεις προσιτές και σε μικρομεσαίες επιχειρήσεις είναι η δημιουργία της Cloud τεχνολογίας, όπου πλέον δεν είναι απαραίτητη η ύπαρξη ενός συστήματος μέσα στην επιχείρηση για την επεξεργασία των δεδομένων, αλλά με μια συνδρομή δίνεται η δυνατότητα αποθήκευσης δεδομένων σε έναν server όχι τοπικά αλλά διαδικτυακά. Τη δυνατότητα αυτή την προσφέρουν ειδικές υπηρεσίες που διαθέτουν την υποδομή, την τεχνογνωσία αλλά είναι παράλληλα και σε θέση να προσφέρουν τις απαραίτητες εγγυήσεις για την ασφάλεια των πληροφοριών της επιχείρησης.

Κατηγορίες Μεθόδων Πρόβλεψης

Οι μέθοδοι πρόβλεψης μπορούν να διαχωριστούν σε δυο βασικές κατηγορίες, τις ποσοτικές και τις κριτικές προβλέψεις.

2.5

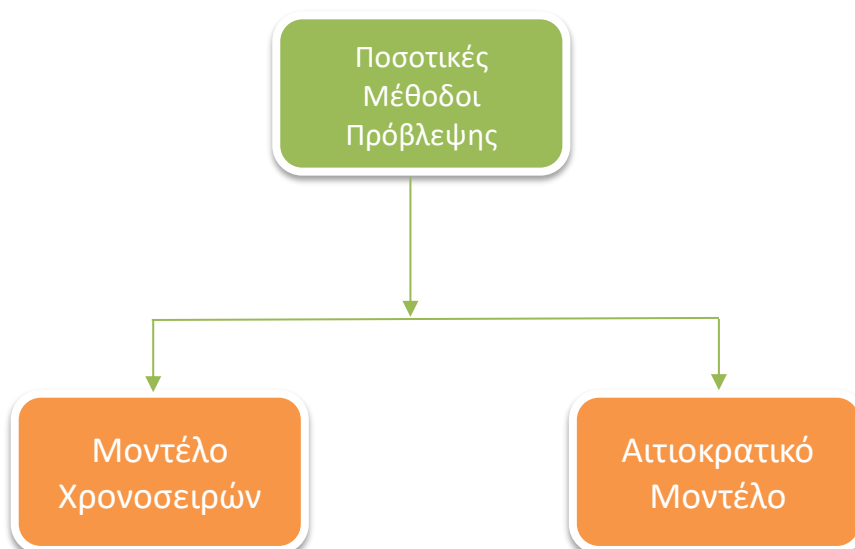


Εικόνα 2.4 Κατηγορίες μεθόδων πρόβλεψης

Οι κριτικές προβλέψεις βασίζονται στη γνώση και την εμπειρία ανθρώπων που ασχολούνται με το αντικείμενο προς μελέτη και εμπειρογνομόνων. Χρησιμοποιούνται εν γένει για να παρέχουν συμβουλές και βοήθεια στο σχεδιασμό της στρατηγικής μια επιχείρησης μέσω της ερμηνείας των δεδομένων ως απόρροια πολύχρονης επαγγελματικής εμπειρίας. Παρά το γεγονός ότι η εγκυρότητα τους έχει πολλές φορές αμφισβητηθεί λόγω της υποκειμενικότητας που εμπεριέχεται, συχνά παρέχουν πολύ σημαντικές πληροφορίες στα ανώτερα στελέχη για τη λήψη αποφάσεων. Σε πολλές περιπτώσεις, οι κριτικές προβλέψεις αποτελούν και τη μόνη επιλογή, όπως όταν υπάρχει έλλειψη ιστορικών δεδομένων ή όταν εισάγεται ένα νέο προϊόν στην αγορά. Μια άλλη περίπτωση, είναι όταν ανακύπτει μια κρίση στην αγορά λόγω εμφάνισης μοναδικών συνθηκών, όπως τώρα λόγου χάρη, με την κρίση που έχει επέλθει λόγω της παγκόσμιας πανδημίας του κορονοϊού.

Οι ποσοτικές μέθοδοι πρόβλεψης, βασίζονται σε μαθηματικά μοντέλα ή μοντέλα μηχανικής μάθησης τα οποία χρησιμοποιούν τα διαθέσιμα ιστορικά δεδομένα. Για την εφαρμογή των ποσοτικών μεθόδων πρόβλεψης υπάρχει η απαίτηση η πληροφορία που θέλουμε να προβλέψουμε να ποσοτικοποιείται με την μορφή αριθμητικών δεδομένων και τα ιστορικά δεδομένα να διατηρούν το πρότυπο συμπεριφοράς τους στο μέλλον.

Οι ποσοτικές μέθοδοι πρόβλεψης, μπορούν να χρησιμοποιηθούν ξεχωριστά είτε και σε συνδυασμό με τις κριτικές προβλέψεις.



Εικόνα 2.5 Κατηγοριοποίηση των ποσοτικών μεθόδων πρόβλεψης

Τα μοντέλα που αφορούν στις ποσοτικές μεθόδους πρόβλεψης διακρίνονται σε δυο μεγάλες κατηγορίες: το μοντέλο χρονοσειρών και το αιτιοκρατικό ή επεξηγηματικό μοντέλο. Το αιτιοκρατικό μοντέλο στηρίζεται στη βασική υπόθεση ότι υπάρχει μια σταθερή σχέση μεταξύ του υπό πρόβλεψη μεγέθους (εξαρτημένη μεταβλητή) και ορισμένων παραμέτρων (ανεξάρτητες μεταβλητές) που το επηρεάζουν. Οπότε, για την εφαρμογή του συγκεκριμένου μοντέλου, απαραίτητη προϋπόθεση είναι, η ύπαρξη ιστορικών δεδομένων και για το υπό εξέταση μέγεθος αλλά και για κάποιες άλλες ανεξάρτητες μεταβλητές. Ένα χαρακτηριστικό παράδειγμα αιτιοκρατικού μοντέλου είναι, η πρόβλεψη για κατανάλωση ενέργειας κατοικίας, όπου οι καταναλωτές εμφανίζουν κάποια συγκεκριμένα μοτίβα κατανάλωσης αλλά συγχρόνως κάποιες μεταβλητές (ανεξάρτητες) όπως είναι η θερμοκρασία, επηρεάζουν σημαντικά την κατανάλωση ενέργειας.

Το μοντέλο χρονοσειρών θα αναλυθεί εκτενώς στην επόμενη ενότητα.

Χρονοσειρά

Εισαγωγή

2.6 Η χρονοσειρά είναι ένα σύνολο αριθμητικών παρατηρήσεων που καταγράφονται για μια χρονική περίοδο. Είναι μια ακολουθία δεδομένων που παρατηρείται συνήθως σε τακτά χρονικά διαστήματα. Η ανάλυση χρονοσειρών προβλέπει την συμπεριφορά ενός φυσικού ή άλλου μεγέθους μελετώντας τις διαδοχικές παρατηρήσεις αυτού στο παρελθόν. Εάν οι παρατηρήσεις είναι μεταξύ τους εξαρτημένες, οι μελλοντικές τιμές μπορούν να προσδιοριστούν ακριβώς από τις προηγούμενες και η διαδικασία ονομάζεται ντετερμινιστική. Στην αντίθετη περίπτωση όπου, οι μελλοντικές τιμές καθορίζονται μερικώς από τις τιμές του παρελθόντος, οι χρονοσειρές θεωρείται ότι αντιπροσωπεύουν στοχαστικές διαδικασίες. Στον πραγματικό κόσμο και ειδικότερα στον κόσμο των επιχειρήσεων, παρατηρούμε μόνο στοχαστικές χρονοσειρές αφού όλα τα φυσικά μεγέθη χαρακτηρίζονται από τον παράγοντα της τύχης και τα μοντέλα που περιέχουν τον τυχαίο παράγοντα ονομάζονται στοχαστικά. Μια χρονοσειρά μπορεί να είναι συνεχής ή διακριτή. Σε μια συνεχή χρονοσειρά, υπάρχουν παρατηρήσεις για κάθε μονάδα του χρόνου ενώ αν η καταγραφή τιμών γίνεται μόνο σε ορισμένες χρονικές στιγμές έχουμε διακριτή χρονοσειρά. Συνήθως, σε μια διακριτή χρονοσειρά καταγράφονται οι διαδοχικές παρατηρήσεις σε ίσα χρονικά διαστήματα όπως ωριαία, ημερήσια, εβδομαδιαία, μηνιαία και ετήσια.

2.6.2

Μοντέλο Χρονοσειρών

Το μοντέλο χρονοσειρών αποτελεί το πιο διαδεδομένο είδος ποσοτικού μοντέλου πρόβλεψης και χρησιμοποιεί για την παραγωγή προβλέψεων δεδομένα τα οποία αφορούν μόνο το υπό πρόβλεψη μέγεθος. Η μοντελοποίηση χρονοσειρών, είναι ένας δυναμικός τομέας που έχει προσελκύσει την προσοχή της ερευνητικής κοινότητας τις τελευταίες δεκαετίες. Ο κύριος στόχος της μοντελοποίησης χρονοσειρών είναι η προσεκτική συλλογή και σχολαστική μελέτη των προηγούμενων παρατηρήσεων μιας χρονοσειράς με σκοπό αναπτυχθεί το κατάλληλο πρότυπο που περιγράφει καλύτερη την εγγενή δομή της χρονοσειράς. Το εν λόγω πρότυπο προεκτείνεται στο μέλλον και με αυτόν το τρόπο πραγματοποιείται η πρόβλεψη. Η πρόβλεψη χρονοσειρών, επομένως, μπορεί να οριστεί ως η πράξη εκτίμησης του μέλλοντος κατανοώντας το παρελθόν. Είναι σημαντικό να τονίσουμε ότι η εφαρμογή του μοντέλου των χρονοσειρών στηρίζεται στην παραδοχή ότι το εν λόγω πρότυπο αναγνωρίζεται μονοσήμαντα με βάση τα ιστορικά δεδομένα και θεωρείται ότι θα διατηρηθεί σταθερό.

Με αυτό τον τρόπο, το συγκεκριμένο μοντέλο πρόβλεψης είναι κατάλληλο για να υπολογιστεί το μελλοντικό ύψος των πωλήσεων ενός προϊόντος, του οποίου η ποιότητα, η τιμή και το διαφημιστικό κόστος πρόκειται να παραμείνουν σταθερά. Λόγω της αδιαμφισβήτητης σημασίας της πρόβλεψης χρονοσειρών σε μια πληθώρα πρακτικών τομέων, θα πρέπει να δοθεί ιδιαίτερη βαρύτητα στην αναγνώριση του μοντέλου που θα ταιριάζει απόλυτα με την υποκείμενη χρονοσειρά. Χωρίς αμφιβολία, μια επιτυχημένη πρόβλεψη χρονοσειράς εξαρτάται άμεσα με την επιλογή του κατάλληλου μοντέλου, γεγονός που αποτελεί αντικείμενο μελέτης εδώ και πολλά χρόνια με σκοπό τη βελτίωση της ακρίβειας των προβλέψεων. Ως αποτέλεσμα, στη βιβλιογραφία, έχουν αναπτυχθεί διάφορα σημαντικά μοντέλα πρόβλεψης.

Ανάλυση Χρονοσειράς

2.6.3 Η ανάλυση χρονοσειράς αποσκοπεί στην ανεύρεση των χαρακτηριστικών εκείνων που συμβάλουν στην κατανόηση της ιστορικής συμπεριφοράς μιας μεταβλητής και επιτρέπουν την πρόβλεψη μελλοντικών τιμών της. Ο κυριότερος στόχος στην ανάλυση χρονοσειρών είναι η επιλογή και προσαρμογή του κατάλληλου μοντέλου που να προσεγγίζει ικανοποιητικά τα δεδομένα και να περιγράφει το μηχανισμό της χρονοσειράς από την οποία προέκυψε η συγκεκριμένη σειρά, καθώς και η χρησιμοποίηση του μοντέλου για πρόβλεψη. Η μεγαλύτερη πρόκληση στην ανάλυση χρονοσειρών είναι η πρόβλεψη, δηλαδή πως η ακολουθία των παρατηρήσεων θα συνεχιστεί στο μέλλον. Το ζητούμενο είναι να ακολουθεί μια διαδικασία που θα εξασφαλίσει ότι θα παραχθούν όσο το δυνατόν πιο ακριβείς προβλέψεις, αξιοποιώντας στο έπακρο όλη την διαθέσιμη ιστορική πληροφορία

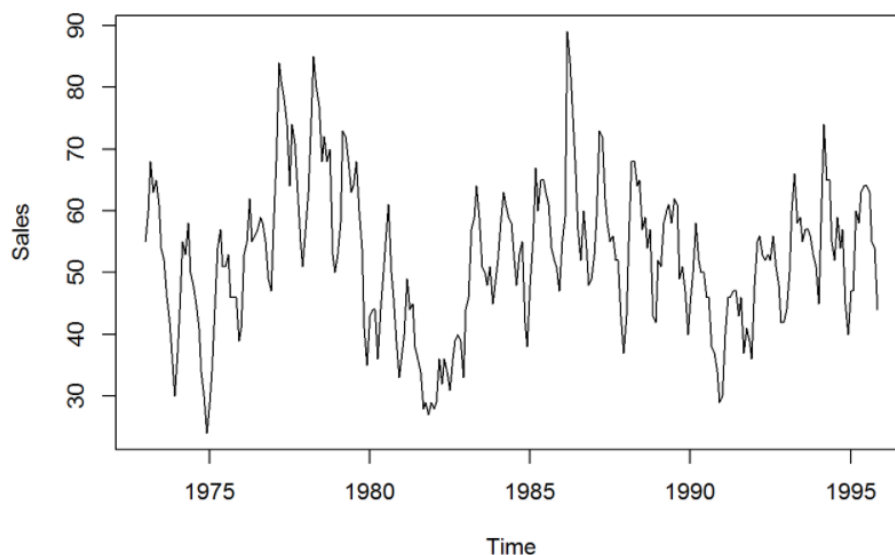
2.6.3.1 Γραφική Αναπαράσταση Χρονοσειράς

Η πρώτη και κύρια διαδικασία που πρέπει να πραγματοποιηθεί κατά την ανάλυση της χρονοσειράς είναι η γραφική αναπαράσταση των δεδομένων της, η δισδιάστατη δηλαδή γραφική απεικόνιση των πραγματικών τιμών των διαθέσιμων δεδομένων ως προς το χρόνο. Από το απλούστερο έως το πιο σύνθετο γράφημα, όλα τα γραφήματα προσφέρουν μια υψηλή αξία τόσο κατά την έναρξη της ανάλυσης της χρονοσειράς όσο και στη λήξη της διαδικασίας όπου επικοινωνούνται τα αποτελέσματα. Μέσω της οπτικοποίησης των δεδομένων που προσφέρουν τα γραφήματα, είναι δυνατόν να διαπιστωθούν μοτίβα, «ιδιάζουσες» τιμές (outliers) - τιμές δηλαδή που βρίσκονται σε προφανή απόκλιση από τις υπόλοιπες καθώς και εσφαλμένες τιμές. Οι τιμές αυτές ενδέχεται να δημιουργήσουν σοβαρά προβλήματα στην μμοντελοποίηση μμιας

χρονοσειράς και ως εκ τούτου χρειάζονται ειδική μμεταχείριση αφού όμως πρώτα προσδιοριστεί το αίτιο το οποίο τις προκάλεσε.

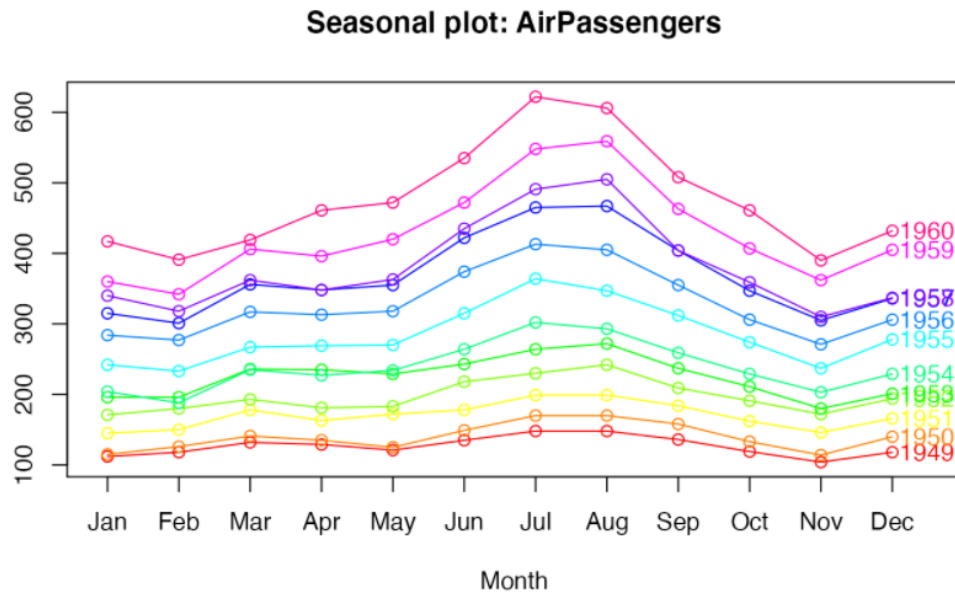
Οι κύριοι τύποι γραφημάτων που χρησιμοποιούνται για την γραφική αναπαράσταση χρονοσειρών είναι:

- **Διαγράμματα Χρόνου (time plots):** Πρόκειται για το πιο διαδεδομένο γράφημα λόγω της ευκολίας ως προς τη χρήση και την κατανόησή του. Οι παρατηρήσεις που ενώνονται με ευθείες γραμμές αναπαριστώνται στην πάροδο του χρόνου. Μέσω αυτών των διαγραμμάτων γίνονται άμεσα αντιληπτά η τάση των δεδομένων, η εποχικότητα όπως και άλλα βασικά συστατικά των χρονοσειρών.



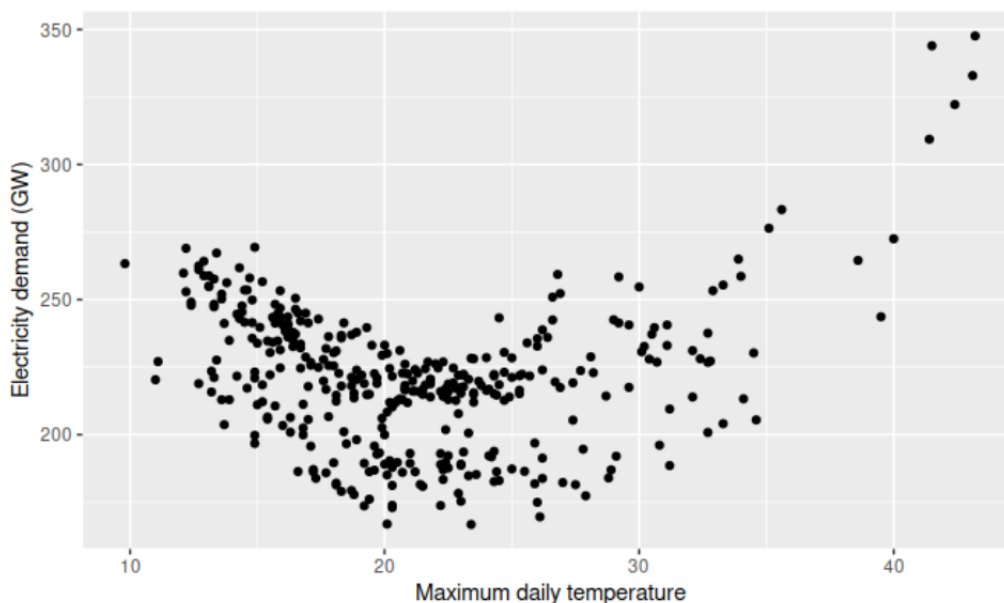
Εικόνα 2.6 Μηνιαίες πωλήσεις σπιτιών για μονογονεϊκές οικογένειες στις ΗΠΑ

- **Εποχιακά Διαγράμματα (seasonal plots):** Ένα εποχιακό διάγραμμα είναι παρόμοιο με το διάγραμμα χρόνου, με τη διαφορά ότι τα δεδομένα απεικονίζονται σε σχέση με τις μεμονωμένες περιόδους που παρατηρούνται τα δεδομένα. Ενδείκνυνται για χρονοσειρές που εμφανίζουν έντονη εποχικότητα και ο αναλυτής μπορεί να παρατηρήσει άμεσα πότε εν δυνάμει αλλάζει το μοτίβο.



Εικόνα 2.7 Αριθμός επιβατών αυστραλιανής αεροπορικής εταιρείας

- **Διαγράμματα Διασποράς (scatter plots) :** Σε αυτό το διάγραμμα κάθε σημείο παρουσιάζει διαφορετικό προϊόν, υπηρεσία ή ακόμα και διαφορετικό τύπο του ίδιου του προϊόντος, και το συνολικό διάγραμμα δείχνει τη σχέση μεταξύ δύο διαφορετικών μεγεθών, χαρακτηριστικών που αφορούν το αντίστοιχο σημείο, τα οποία αναπαρίστανται ποσοτικοποιημένα στους δύο κάθετους άξονες. Μία τέτοιας κατηγορίας γράφημα, βοηθάει στην οπτικοποίηση της σχέσης των μεγεθών για το υπό μελέτη δείγμα και υποδεικνύει αντίστοιχα τις μεταβλητές που πρέπει να παραμετροποιηθούν κατάλληλα στο μοντέλο πρόβλεψης που θα χρησιμοποιηθεί.



Εικόνα 2.8 Ημερήσια ζήτηση ηλεκτρικής ενέργειας

2.6.3.2 Ποιοτικά Χαρακτηριστικά της Χρονοσειράς

Η επισκόπηση του γραφήματος στο πεδίο του χρόνου που αναφέρθηκε παραπάνω, έχει ως αποτέλεσμα την ανίχνευση των ποιοτικών χαρακτηριστικών της χρονοσειράς που είναι: η τάση, η κυκλικότητα, η εποχιακότητα και οι μη κανονικές διακυμάνσεις.



Εικόνα 2.9 Ποιοτικά χαρακτηριστικά της χρονοσειράς

Η τάση, αφορά στη συστηματική αύξηση ή μείωση της τιμής του μεγέθους υπό μελέτη. Μπορεί να είναι γραμμική, δηλαδή να εξελίσσεται στη μορφή μιας ευθείας ή στη μορφή οποιασδήποτε καμπύλης. Ανάλογα την περίπτωση, ο αναλυτής πρέπει να εκτιμήσει το κατάλληλο μήκος περιόδου για την αναζήτηση ύπαρξης τάσης.

Το δεύτερο και σημαντικότερο χαρακτηριστικό για την απόκτηση ενός μοντέλου με καλή ακρίβεια είναι η εποχιακότητα, η οποία ορίζεται ως επανάληψη κάποιου είδους μοτίβου, που εμφανίζεται με πανομοιότυπο τρόπο σε ίδια χρονικά διαστήματα όπως

μέρες, εβδομάδες, μήνες, κ.λπ. Χαρακτηριστικό παράδειγμα εποχιακής χρονοσειράς αποτελούν οι πωλήσεις των χριστουγεννιάτικων ειδών όπου εμφανίζουν έντονη άνοδο τον Νοέμβριο και τον Δεκέμβριο. Στο σημείο αυτό, αξίζει να επισημανθεί ότι εάν κάποια χρονοσειρά δεν έχει ούτε εποχιακότητα ούτε τάση, είναι δηλαδή σταθερού επιπέδου, τότε χαρακτηρίζεται από στασιμότητα.

Η κυκλικότητα, ομοίως με την εποχιακότητα παρουσιάζει μια μεταβολή που εμφανίζεται κατά περιόδους. Η ειδοποιός διαφορά, έγκειται στο γεγονός ότι στην κυκλικότητα το μοτίβο ξαναεμφανίζεται μετά από κάποια χρονική περίοδο, όχι απαραίτητα σταθερή αλλά και η ένταση μπορεί να διαφοροποιείται συγκριτικά με την προηγούμενη περίοδο εμφάνισης. Συνήθως οφείλεται σε εξωγενείς συνθήκες και κυρίως εμφανίζεται σε βασικά οικονομικά μεγέθη, όπως, για παράδειγμα, μια οικονομική ύφεση η οποία έχει διαφορετικό αντίκτυπο συγκριτικά με τις προηγούμενες υφέσεις και το βήμα του κύκλου της δεν είναι αυστηρά προκαθορισμένο.

Μη κανονικές διακυμάνσεις θεωρούνται εκείνες που μπορεί να αντιπροσωπεύουν μια εντελώς τυχαία μεταβλητή ή μπορεί να αντιπροσωπεύουν κάποια ασυνέχεια που συνδέεται με κάποιο εξαιρετικό γεγονός. Οι ασυνέχειες, θεωρούνται υποσύνολο των μη κανονικών διακυμάνσεων, και ουσιαστικά αποτελούν κάποιες ακραίες τιμές ή κάποια αλλαγή επιπέδου. Οι ακραίες τιμές, ενώ οι επίδρασή τους στην χρονοσειρά έχει μικρή χρονική διάρκεια χρήζουν ιδιαίτερη προσοχή καθώς είναι εκείνες που προκαλούν πρόβλημα στην εφαρμογή κάποιου προτύπου συμπεριφοράς έτσι ώστε να προκύψει μια μελλοντική τιμή για την αντίστοιχη χρονοσειρά. Η αλλαγή επιπέδου παρουσιάζει πιο μόνιμο χαρακτήρα και μπορεί να προκληθεί, για παράδειγμα, σε περιπτώσεις όπου εισέρχεται ένας ανταγωνιστής στην αγορά ενός προϊόντος και παίρνει ένα μερίδιο της αγοράς. Μη κανονικές διακυμάνσεις εμφανίζονται πάντα στις χρονοσειρές.

Βασικά στάδια στη διαδικασία πρόβλεψης

Η διαδικασία της πρόβλεψης περιλαμβάνει πέντε βασικά βήματα τα οποία θα αναλυθούν παρακάτω:

2.7 Καθορισμός Προβλήματος (Problem Definition)

Πρώτο βήμα στη διαδικασία πρόβλεψης και αρκετές φορές και το πιο δύσκολο αποτελεί ο προσδιορισμός του αντικειμένου προς πρόβλεψη. Πρόκειται για ένα στάδιο που απαιτεί κατανόηση σε βάθος του τρόπου που θα χρησιμοποιηθεί η παραχθείσα πρόβλεψη, των ατόμων που θα χρειαστούν την πρόβλεψη και γενικά ο ρόλος της πρόβλεψης για τον εμπλεκόμενο οργανισμό, επιχείρηση, κ.λπ.

Συγκέντρωση Πληροφοριών (Gathering Information)

Η ορθή συλλογή δεδομένων μπορεί να είναι σε ορισμένες περιπτώσεις μια δύσκολη και χρονοβόρα διαδικασία και για αυτό θα πρέπει να της δίνεται ιδιαίτερη βαρύτητα, ώστε να αντιμετωπίζονται πιθανά προβλήματα που σχετίζονται με τη διαθεσιμότητά τους, την ακρίβεια τους κ.α. Σε αυτό το βήμα απαιτούνται τουλάχιστον δύο είδη πληροφοριών. Το πρώτο είναι τα στατιστικά (συνήθως αριθμητικά) δεδομένα και το δεύτερο η κρίση, η πείρα και η εμπειρία του προσωπικού που ασχολούνταν με αυτή τη συλλογή για αυτό το χρονικό διάστημα.

Προκαταρκτική Ανάλυση (Exploratory Analysis)

Στο βήμα αυτό, μας απασχολεί το είδος της πληροφορίας που αποκομίζουμε από τα ακατέργαστα ιστορικά δεδομένα. Αρχικά, αναπαριστούμε γραφικά τα δεδομένα με σκοπό να αποκτήσουμε μία αίσθηση των δεδομένων, δίνοντας απαντήσεις σε ερωτήματα όπως αν υπάρχουν λανθασμένα πρότυπα, αν υπάρχει σημαντική τάση ή εποχικότητα και τέλος, αν υπάρχουν ασυνήθιστες τιμές (outliers). Χρήσιμο εργαλείο στο στάδιο αυτό αποτελεί η διαδικασία της αποσύνθεσης που απομονώνει τις 4 βασικές συνιστώσες που αναφέρθηκαν παραπάνω. Σκοπός της ανάλυσης είναι η προετοιμασία των χρονοσειρών για την εφαρμογή μοντέλων πρόβλεψης που αναμένεται να δώσουν ικανοποιητικές προβλέψεις.

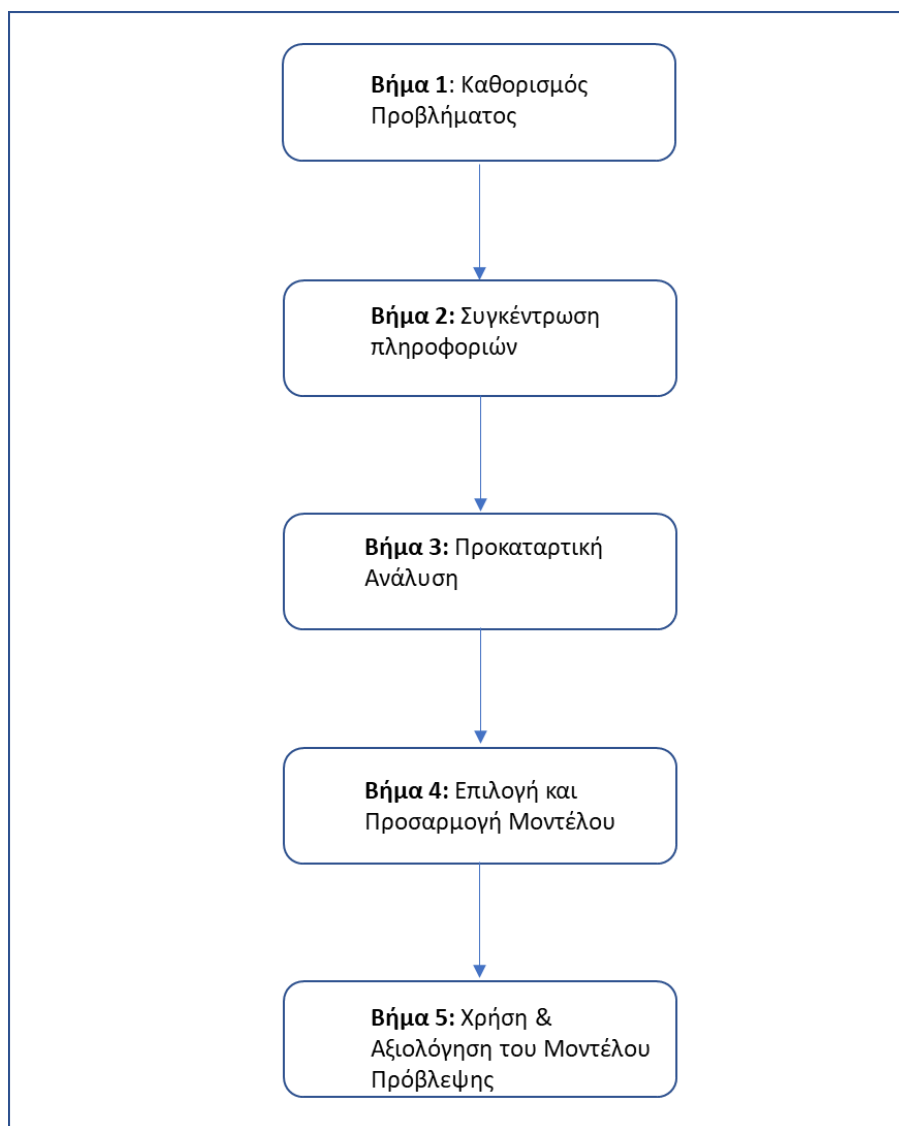
Επιλογή και Προσαρμογή Μοντέλου (Choosing and Fitting Models)

Η παραγωγή προβλέψεων λαμβάνει χώρα με την στατιστική ή κριτική προέκταση των διαθέσιμων δεδομένων. Ο αναλυτής καλείται στο βήμα αυτό να επιλέξει τα κατάλληλα

μοντέλα πρόβλεψης αλλά και τις παραμέτρους τους, βάσει των ιστορικών δεδομένων αλλά και των ιδιοτήτων της κάθε μεθόδου.

Χρήση και Αξιολόγηση του Μοντέλου Πρόβλεψης (Using and Evaluating a Forecasting Model)

Εφόσον το μοντέλο έχει επιλεγεί και οι παράμετροί του έχουν καθοριστεί, στο τελικό στάδιο χρησιμοποιείται ώστε να παραχθούν οι ζητούμενες προβλέψεις. Η απόδοση του μοντέλου μπορεί να αξιολογηθεί σωστά με την πάροδο του χρόνου, καθώς νέα δεδομένα γίνονται διαθέσιμα. Θα πρέπει να εξεταστεί μια γκάμα κατάλληλων στατιστικών δεικτών για την αξιολόγηση και μέτρηση της ακρίβειας των προβλέψεων.



Εικόνα 2.10 Τα βήματα της διαδικασίας πρόβλεψης

Παράγοντες που επηρεάζουν την επιλογή μεθόδου πρόβλεψης

2.8 Μια εταιρεία χρειάζεται να λαμβάνει υπόψιν ουκ ολίγους παράγοντες προκειμένου να επιλέξει τη μέθοδο πρόβλεψης που θα χρησιμοποιήσει, συμπεριλαμβανομένης της ζήτησης των προηγούμενων χρόνων, των ενεργειών του τμήματος marketing, της στρατηγική που έχει χαράξει, της ευρύτερη κατάσταση της οικονομίας, ακόμα και των ενεργειών από πλευράς ανταγωνιστών.

Σήμερα οι μέθοδοι που έχουν αναπτυχθεί για τη διενέργεια των προβλέψεων είναι αρκετές και χρησιμοποιούνται για τη λήψη των κατάλληλων αποφάσεων ανάλογα πάντα με τις εκάστοτε συνθήκες της επιχείρησης. Παρακάτω παρουσιάζονται κάποια κριτήρια που τίθενται υπό αξιολόγηση κατά τη διαδικασία επιλογής της κατάλληλης μεθόδου ζήτησης.

Η περίοδος και ο ορίζοντας πρόβλεψης: Περίοδος πρόβλεψης είναι το χρονικό διάστημα, με βάση την οποία γίνονται οι προβλέψεις. Περίοδος πρόβλεψης μπορεί να είναι η εβδομάδα ή ο μήνας ή το έτος, στην περίπτωση που μελετάμε είναι ο μήνας. Ορίζοντας πρόβλεψης είναι ένας αριθμός περιόδων, για τις οποίες ζητείται να γίνει πρόβλεψη της ζήτησης. Τόσο η περίοδος όσο και ο ορίζοντας πρόβλεψης υπαγορεύονται από το είδος της απόφασης (στρατηγική, τακτική, λειτουργική), ενώ ανάλογα υπαγορεύεται και η μέθοδος πρόβλεψης που χρησιμοποιείται. Για παράδειγμα, για στρατηγικές αποφάσεις (μακροπρόθεσμες αποφάσεις), ο ορίζοντας πρόβλεψης μπορεί να είναι 5 έτη και ως περίοδος πρόβλεψης να ορίζεται το έτος.

Τα διαθέσιμα στοιχεία: Το είδος και η ποσότητα των διαθέσιμων στοιχείων επηρεάζουν την επιλογή της μεθόδου, αφού κάθε μέθοδος έχει διαφορετικές απαιτήσεις. Ενδεικτικά, η εμφάνιση εποχικότητας στα δεδομένα επηρεάζει την επιλογή της μεθόδου πρόβλεψης. Επίσης, για την εφαρμογή αλγορίθμων μηχανικής μάθησης είναι απαραίτητος ένας μεγάλος όγκος δεδομένων.

Το κόστος της μεθόδου: Το κόστος αυτό διαμορφώνεται σε διαφορετικό ύψος, ανάλογα με την επιζητούμενη ακρίβεια πρόβλεψης και τις απαιτήσεις που έχει η μέθοδος όσον αφορά το πλήθος των στοιχείων, τις αναγκαίες επεξεργασίες, το χρόνο απασχόλησης όσων εμπλακούν στις διαδικασίες ανάπτυξης της μεθόδου, προετοιμασίας των στοιχείων, εγκατάστασης και λειτουργίας της μεθόδου κλπ. Το

κόστος δεν αφορά μόνο τη χρήση εξειδικευμένου λογισμικού(κόστος αδειών, κ.λπ.) αλλά και τη χρήση εξειδικευμένου προσωπικού(κόστος εκπαίδευσης, κ.λπ.).

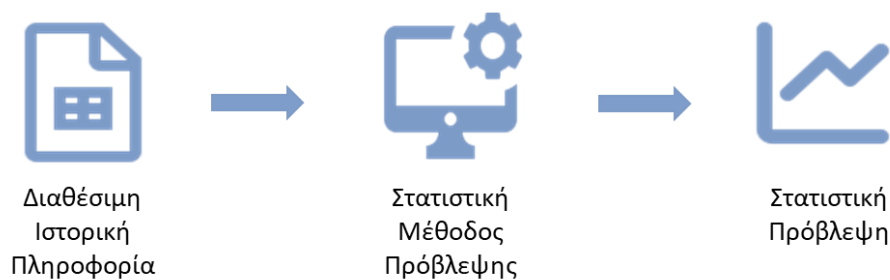
Η επιζητούμενη ακρίβεια: Το επίπεδο ακρίβειας των προβλέψεων είναι μια παράμετρος που καθορίζεται από το χρήστη, ανάλογα με το είδος των προβλέψεων και την ασφάλεια των προβλέψεων (ακρίβεια) που επιδιώκει να πετύχει. Η ακρίβεια μιας μεθόδου αυξάνει όσο περισσότερο στηρίζεται σε ποσοτικά στοιχεία, όσο περισσότερα είναι τα στοιχεία αυτά και όσο μικρότερος είναι ο χρονικός ορίζοντας των προβλέψεων.

Η απλότητα και ευκολία εφαρμογής: Οι απλούστερες μέθοδοι είναι ευκολότερο να κατανοηθούν και να υιοθετηθούν από τη διοίκηση μιας επιχείρησης, ώστε να επικουρούν στη λήψη των αποφάσεων. Η απλότητα συνήθως συνδυάζεται με ευκολότερη εφαρμογή και χαμηλότερο κόστος ανάπτυξης, εγκατάστασης και εφαρμογής της μεθόδου.

Μοντέλα πρόβλεψης

Στατιστικά Μοντέλα

- Ο κυριότερος στόχος στην ανάλυση χρονοσειρών είναι η επιλογή και προσαρμογή του κατάλληλου μοντέλου που να προσεγγίζει ικανοποιητικά τα δεδομένα και να περιγράφει το μηχανισμό της χρονοσειράς από την οποία προέκυψε η συγκεκριμένη σειρά, καθώς και η χρησιμοποίηση του μοντέλου για πρόβλεψη.



Εικόνα 3.1 Ροή διαδικασίας της Στατιστικής Πρόβλεψης

Στο παρόν κεφάλαιο παρουσιάζονται συνοπτικά οι βασικότερες στατιστικές μέθοδοι πρόβλεψης που απασχολούν την επιστήμη των προβλέψεων. Στατιστική πρόβλεψη είναι η διαδικασία κατά την οποία εισάγεται η διαθέσιμη ιστορική πληροφορία-δεδομένα, υπάρχει μια μέθοδος που επεξεργάζεται αυτά τα δεδομένα και το αποτέλεσμα που εξάγεται είναι η τελική πρόβλεψη. Στη βιβλιογραφία ορίζεται ως εφαρμογή στατιστικών μοντέλων χρονοσειρών ή αιτιοκρατικών μοντέλων επί μιας σειράς δεδομένων με σκοπό την αυτοματοποιημένη και συστηματική παραγωγή προβλέψεων.

Τα κύρια πλεονεκτήματα των στατιστικών προβλέψεων είναι:

- Δεν προϋποθέτουν εξειδικευμένες γνώσεις
- Άμεσα εφαρμόσιμες και αποδεκτά ακριβείς (αν συνδυαστούν με κατάλληλα διαστήματα εμπιστοσύνης)
- Η εφαρμογή τους απαιτεί ελάχιστο χρόνο
- Απαιτούν λίγους υπολογιστικούς πόρους

Οι στατιστικές προβλέψεις, όμως, έχουν και κάποια αξιοσημείωτα μειονεκτήματα που αναφέρονται παρακάτω:

- Προϋποθέτουν ότι το μοτίβο που ακολουθούν τα τρέχοντα δεδομένα θα συνεχιστεί και στο μέλλον, μια συνθήκη που δεν ικανοποιείται πάντα
- Δεν λαμβάνουν υπόψη ειδικά γεγονότα και ενέργειες που ενδέχεται να πραγματοποιηθούν (πχ. διαφημίσεις, αθλητικά συμβάντα)
- Αρκετές στατιστικές μέθοδοι απαιτούν αρκετές παρατηρήσεις ιστορικών δεδομένων που είναι μπορεί να γίνει μια διαδικασία δύσκολη και χρονοβόρα

Μέθοδος Naïve

3.1.1 Αν και αποτελεί την πιο απλή στατιστική μέθοδο-το υποδεικνύει και η ονομασία της- καθώς δίνει ως πρόβλεψη για την επόμενη χρονική στιγμή την πραγματική τιμή της προηγούμενης χρονικής στιγμής, χρησιμοποιείται κατά κόρον στις στατιστικές προβλέψεις ως σημείο αναφοράς (benchmark) για άλλες, πιο πολύπλοκες μεθόδους. Δε λαμβάνει υπόψη της ούτε την τάση ούτε την εποχιακότητα και σε πολύ συγκεκριμένες περιπτώσεις όπου υπάρχει μεγάλη τυχαιότητα στα δεδομένα, στον τομέα των χρηματοοικονομικών, η συγκεκριμένη πρόβλεψη μπορεί να είναι ανταγωνιστική έναντι πιο περίπλοκων μεθόδων.

3.1.2 Seasonal Naïve

Πρόκειται για μια παραλλαγή της προηγούμενης μεθόδου και εφαρμόζεται σε δεδομένα με έντονη εποχιακότητα. Η διαφορά έγκειται στο γεγονός ότι η πρόβλεψη είναι ίση με την τελευταία χρονική στιγμή της αντίστοιχης εποχιακής περιόδου. Σε 3.1.3 χρονοσειρές με τάση δεν έχει ιδιαίτερα ικανοποιητική ακρίβεια ενώ παρουσιάζει αξιόλογα αποτελέσματα σε χρονοσειρές σταθερού επιπέδου με έντονη εποχιακότητα.

Μοντέλα Παλινδρόμησης

Μέσω των μοντέλων αυτών, εξετάζεται η σχέση μεταξύ μιας εξαρτημένης μεταβλητής με συγκεκριμένες ανεξάρτητες μεταβλητές. Τα μοντέλα αυτά διακρίνονται σε απλά και πολλαπλά, ανάλογα με το πλήθος των ανεξάρτητων μεταβλητών. Τα μοντέλα αυτά είναι εύκολα στη χρήση και θεωρούνται αρκετά αξιόπιστα για μακροπρόθεσμες προβλέψεις.

Μέθοδοι Εκθετικής Εξομάλυνσης

Πρόκειται για μια οικογένεια μοντέλων που έγιναν ιδιαίτερα δημοφιλή λόγω της ευκολίας, της ελάχιστης απαίτησης σε υπολογιστικό χρόνο και σε ιστορικά δεδομένα προκειμένου να παραχθούν προβλέψεις. Είναι από τα πιο παλιά στατιστικά μοντέλα 3.1.4 και είναι κατάλληλα κυρίως για βραχυπρόθεσμες και μεσοπρόθεσμες προβλέψεις ενώ η απόδοσή τους είναι καλύτερη σε δεδομένα που παρουσιάζουν στασιμότητα ή μικρό ρυθμό ανάπτυξης. Από τις πιο γνωστές μεθόδους της εν λόγω κατηγορίας είναι η απλή εκθετική εξομάλυνση σταθερού επιπέδου, η εκθετική εξομάλυνση γραμμικής τάσης και η εκθετική εξομάλυνση μη γραμμικής ή φθίνουσας τάσης. Χρησιμοποιούνται σε ένα μεγάλο εύρος εφαρμογών από την διαχείριση αποθηκών μέχρι τον τομέα της ηλεκτρικής ενέργειας. Στην παρούσα διπλωματική χρησιμοποιήθηκε ο αλγόριθμος εκθετικής εξομάλυνσης ETS το οποίο αποτελεί μια ευρεία οικογένεια των μεθόδων της εκθετικής εξομάλυνσης και θα αναλυθεί εκτενέστερα στη συνέχεια.

3.1.4.1 Αλγόριθμος Εκθετικής Εξομάλυνσης ETS

Πρόκειται για μια σειρά μοντέλων state-space, το καθένα από τα οποία αποτελείται από μια εξίσωση που περιγράφει τα παρατηρούμενα δεδομένα και κάποιες εξισώσεις κατάστασης που περιγράφουν χαρακτηριστικά που αλλάζουν με τον χρόνο (level,trend,seasonal). Το όνομα ETS είναι αρχικά για τα Error, Trend, Seasonal τα οποία είναι τα στοιχεία που χρησιμοποιεί η μέθοδος για την παραγωγή προβλέψεων. Για κάθε μέθοδο υπάρχουν δύο μοντέλα, ένα με προσθετικά σφάλματα και ένα με πολλαπλασιαστικά. Αντίστοιχα, η εποχιακότητα μπορεί να είναι προσθετική πολλαπλασιαστική ή να μην λαμβάνεται καθόλου υπόψη. Τέλος όσον αφορά την τάση μπορεί να μην λαμβάνεται καθόλου υπόψη(None), μπορεί να είναι προσθετική(Additive), προσθετική φθίνουσα(Additive Damped), πολλαπλασιαστική(Multiplicative) και πολλαπλασιαστική φθίνουσα(Multiplicative Damped). Οι αντίστοιχοι συμβολισμοί είναι:

- ο Σφάλματα: Error = {A,M}
- ο Εποχιακότητα: Seasonal = {N,A,M}
- ο Τάση: Trend = {N,A,Ad,M,Md}

Συνεπώς υπάρχουν 30 μοντέλα state-space.

Κάθε μοντέλο έχει ένα στοιχείο τάσης και ένα εποχιακό στοιχείο όπως δίνεται στον παρακάτω πίνακα

	Seasonal Component		
Trend	N	A	M
Component	(None)	(Additive)	(Multiplicative)
N (None)	(N,N)	(N,A)	(N,M)
A (Additive)	(A,N)	(A,A)	(A,M)
A _d (Additive damped)	(A _d ,N)	(A _d ,A)	(A _d ,M)
M (Multiplicative)	(M,N)	(M,A)	(M,M)
M _d (Multiplicative damped)	(M _d ,N)	(M _d ,A)	(M _d ,M)

Εικόνα 3.2 Ταξινόμηση των μοντέλων εκθετικής εξομάλυνσης

Η πραγματική τιμή μπορεί να γραφτεί ως η τιμή του προηγούμενου επιπέδου συν ένα σφάλμα, οπότε προκύπτει η εξίσωση:

$$y_t = \ell_{t-1} + e_t$$

Για ένα state space model αρκεί να προσδιοριστεί η κατανομή πιθανοτήτων για e_t . Αν στο μοντέλο με προσθετικά σφάλματα υποτεθεί ότι τα σφάλματα ενός βήματος κατανέμονται κανονικά με μέση τιμή 0 και διακύμανση σ^2 μπορούμε να χρησιμοποιήσουμε την συντομογραφία $e_t = \varepsilon_t \sim \text{NID}(0, \sigma^2)$, όπου NID (Normally and Independently Distributed) κατανεμημένα κανονικά και ανεξάρτητα. Προκύπτουν επομένως οι εξισώσεις:

$$y_t = \ell_{t-1} + \varepsilon_t$$

$$\ell_t = \ell_{t-1} + \alpha \varepsilon_t$$

Η πρώτη εξίσωση ονομάζεται εξίσωση μέτρησης ενώ η δεύτερη εξίσωση κατάστασης ή μετάβασης. Με αυτές τις εξισώσεις και την κατανομή των σφαλμάτων καθορίζεται

πλήρως ένα μοντέλο. Η εξίσωση μέτρησης εκφράζει τη σχέση μεταξύ παρατηρήσεων και μη παρατηρούμενων καταστάσεων. Στην περίπτωση αυτή, η παρατήρηση y_t είναι γραμμική συνάρτηση του επιπέδου ℓ_{t-1} και το τυχαίο σφάλμα ε_t , το απρόβλεπτο μέρος του y_t . Για άλλα μοντέλα αυτή η σχέση μπορεί να είναι μη γραμμική. Η εξίσωση κατάστασης δείχνει την εξέλιξη της κατάστασης στον χρόνο. Όσο υψηλότερη είναι η τιμή του α τόσο γρηγορότερες θα είναι και οι μεταβολές του επιπέδου. Όμοια μπορούμε να φτιάξουμε το μοντέλο με πολλαπλασιαστικά σφάλματα, γράφοντας τα τυχαία σφάλματα σαν σχετικά σφάλματα.

Στην ETS μπορούν να χρησιμοποιηθούν ορισμένα κριτήρια πληροφορίας για τον καθορισμό των καλύτερων μοντέλων για μια συγκεκριμένη χρονοσειρά.

Συγκεκριμένα τα κριτήρια αυτά είναι :

- Το κριτήριο πληροφορίας του Akaike ή αλλιώς Akaike's Information Criteria (AIC) το οποίο ορίζεται ως

$$AIC = -2\log(L) + 2k$$

- Το διορθωμένο κριτήριο του Akaike για μικρή προκατάληψη ή αλλιώς AIC_c ορίζεται ως εξής:

$$AIC_c = AIC + \frac{2k(k+1)}{T-k-1}$$

- Το μπεϋζιανό κριτήριο πληροφορίας (Bayesian Information Criterion -BIC)

$$BIC = AIC + k[\log(T)-2]$$

3.1.5

Μέθοδος Theta

Η συγκεκριμένη μέθοδος έγινε γνωστή το 2000 λόγω της πρώτης θέσης που κατέλαβε στο διεθνή διαγωνισμό προβλέψεων M3. Η μέθοδος βασίζεται στην μεταβολή των τοπικών καμπυλοτήτων μιας χρονοσειράς μέσα από την παράμετρο θ (Theta), η οποία εφαρμόζεται απευθείας (πολλαπλασιαστικά) στις διαφορές δεύτερης τάξης των δεδομένων. Η καινούργια χρονοσειρά που δημιουργείται διατηρεί την μέση τιμή και κλίση (παλινδρόμησης) της αρχικής χρονοσειράς αλλά όχι και τις τοπικές καμπυλότητες. Οι χρονοσειρές που παράγονται με αυτή την διαδικασία ονομάζονται γραμμές Θ (Theta Lines). Βασικό ποιοτικό χαρακτηριστικό αυτών των γραμμών είναι η καλύτερη προσέγγιση της μακροπρόθεσμης συμπεριφοράς-τάσης των δεδομένων ή ανάδειξη-τονισμός των βραχυπρόθεσμων χαρακτηριστικών, ανάλογα με την τιμή της

παραμέτρου θ . Ουσιαστικά, βασίζεται σε μια μεθοδολογία αποσύνθεσης της αρχικής χρονοσειράς όπου αναπτύσσεται σε δύο ή περισσότερες γραμμές Theta εκ των οποίων κάθε μία προεκτείνεται ξεχωριστά και προκύπτουν οι προβλέψεις τους που στο τέλος συνδυάζονται. Η προέκταση των γραμμών Theta μπορεί να γίνει με κάθε μία από τις διαθέσιμες μεθόδους πρόβλεψης ενώ μπορεί να εφαρμοστεί διαφορετικός συνδυασμός γραμμών Theta για διαφορετικό ορίζοντα πρόβλεψης. Η μέθοδος Theta λειτουργεί σαν ένας μεγεθυντικός φακός μέσα από τον οποίο οι διακυμάνσεις της χρονοσειράς μεγεθύνονται ή μικραίνουν. Ο γραμμικός συνδυασμός των προβλέψεων των συνιστωσών, γίνεται, μέσα από αυτήν την διαδικασία, πιο αποδοτικός.

Τα βήματα που περιγράφουν τη μεθοδολογία της κλασσικής μεθόδου Theta είναι τα παρακάτω:

Βήμα 0: *Έλεγχος εποχιακότητας*. Ελέγχεται η κάθε χρονοσειρά για στατιστικά σημαντική εποχιακή συμπεριφορά.

Βήμα 1: *Αποεποχικοποίηση*. Μέσω της κλασσικής μεθόδου πολλαπλασιαστικής αποσύνθεσης, εφόσον αποδειχθεί ότι η χρονοσειρά έχει σημαντική εποχιακότητα.

Βήμα 2: *Αποσύνθεση*. Η κάθε χρονοσειρά αποσυντίθεται σε γραμμές Theta, την ευθεία γραμμικής παλινδρόμησης ($\theta = 0$) και τη γραμμή Theta με παράμετρο $\theta = 2$.

Βήμα 3: *Πρόβλεψη*. Η γραμμή Theta με παράμετρο $\theta = 0$, που αναπαριστά την ευθεία γραμμικής παλινδρόμησης, προεκτείνεται με τον συνηθισμένο τρόπο, ενώ η δεύτερη γραμμή προεκτείνεται μέσω της απλής γραμμικής εξομάλυνσης.

Βήμα 4: *Συνδυασμός*. Οι παραγόμενες προβλέψεις των δύο γραμμών Theta συνδυάζονται με ίσα βάρη.

Βήμα 5: *Εποχικοποίηση*. Οι τελικές προβλέψεις εποχικοποιούνται, χρησιμοποιώντας τους δείκτες εποχιακότητας που υπολογίστηκαν στο βήμα 1.

Στην R, η συνάρτηση `theta()` απαιτεί αποεποχικοποιημένα δεδομένα για να παρουσιάσει ικανοποιητική απόδοση. Αντιθέτως, η συνάρτηση `thetaf()`, που χρησιμοποιήθηκε στην παρούσα διπλωματική, ελέγχει αυτόματα τις χρονοσειρές για ύπαρξη εποχιακότητας και σε περίπτωση που εμφανίζεται το εποχικό στοιχείο εξαλείφεται με πολλαπλασιαστική αποσύνθεση προτού εφαρμοστεί η μέθοδος και μετά την εφαρμογή επαναεποχικοποιείται η χρονοσειρά.

Μοντέλα ARIMA

3.1.6 Τα ολοκληρωμένα αυτοπαλινδρομικά μοντέλα κινητών μέσων όρων ARIMA(Auto Regressive Integrated Moving Average) είναι στοχαστικά μαθηματικά μοντέλα τα οποία χρησιμοποιούνται για την περιγραφή της διαχρονικής εξέλιξης κάποιου φυσικού μεγέθους. Μελετήθηκαν από τους Box & Jenkins (1971) και συχνά συναντώνται στη βιβλιογραφία με την αντίστοιχη ονομασία. Προσεγγίζουν τη λογική των κλασικών μοντέλων παλινδρόμησης (π.χ. LRL) και εκθετικής εξομάλυνσης (π.χ. SES) με την έννοια ότι συσχετίζουν τις μελλοντικές τιμές της χρονοσειράς με παρελθοντικές της ή/και σφάλματα που εντοπίστηκαν. Παρουσιάζουν μεγάλη ακρίβεια σε βραχυπρόθεσμες προβλέψεις και δεν απαιτούν κάποια ιδιαίτερη προ-επεξεργασία όπως αφαίρεση τάσης και εποχικότητας. Το πλεονέκτημα των μοντέλων ARIMA είναι ότι επιλέγουν από ένα εύρος μηχανισμών εκείνον που μπορεί καλύτερα να αποδώσει τη σχέση της κάθε παρατήρησης με τις προηγούμενές της.

Τα μη εποχιακά ARIMA μοντέλα είναι γνωστά σαν **ARIMA(p, d, q)** και σχηματίζονται από τον γραμμικό συνδυασμό των παρακάτω παραγόντων

- p : Παράγοντες αυτοπαλινδρόμησης(AR)
- d : Παράγοντες κινητών μέσων όρων(MA)
- q : Παράγοντας διαφορίσης της χρονοσειράς(I)

Αν το μοντέλο περιλαμβάνει αποκλειστικά παράγοντες αυτοπαλινδρόμησης αναφέρεται ως AR(p), αν περιλαμβάνει αποκλειστικά παράγοντες κινητών μέσων 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 διαφορίση της χρονοσειράς πριν την εφαρμογή ενός μοντέλου ARMA(p,q) και έχει ως στόχο την αφαίρεση της τάσης από τα δεδομένα.

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.

Με τη χρήση του τελεστή ολίσθησης B το μοντέλο παριστάνεται ως:

$$(1-\phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)(1-B)^d(1-B^m)^N y_t = c + (\theta_1 B - \theta_2 B^2 + \dots + \theta_q B^q) e_t$$

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

Στην παρούσα διπλωματική χρησιμοποιήσαμε την συνάρτηση της R, **auto.arima()** η οποία επιστρέφει το καλύτερο μοντέλο ARIMA σύμφωνα με τα κριτήρια αξιολόγησης μοντέλων AIC, AICc ή BIC που αξιολογούν την απόδοση των υποψήφιων μοντέλων συναρτήσει της πολυπλοκότητας τους και επιλέγεται εκείνο που ελαχιστοποιεί τις τιμές τους.

Συνδυασμός Μεθόδων

- 3.1.7 Μια προσέγγιση η οποία είναι αξιοσημείωτη, είναι ο συνδυασμός μεθόδων και αναφέρεται στον συνδυασμό δύο ή περισσότερων μεθόδων με ίσα ή άνισα βάρη για την παραγωγή των τελικών προβλέψεων. Όπως αναλύσαμε παραπάνω, κάθε μοντέλο έχει συγκεκριμένες ιδιότητες και αποδίδει διαφορετικά ανάλογα τις συνθήκες παρουσιάζοντας δυνατά και τρωτά σημεία αντίστοιχα. Ως εκ τούτου, ο κατάλληλος συνδυασμός μοντέλων όταν εφαρμοστεί με τις «σωστές» βαρύτητες έχει αναμφισβήτητη αποτελεσματικότητα. Χρησιμοποιείται ευρέως σε εφαρμογές μηχανικής μάθησης όπου εκπαιδεύονται πολλά νευρωνικά δίκτυα και η τελική πρόβλεψη προκύπτει ως ένας σταθμισμένος συνδυασμός τους.

Μοντέλα μηχανικής μάθησης

3.2 Σε αυτό το κεφάλαιο εξετάζονται κάποια μοντέλα που βασίζονται σε μηχανική μάθηση. Σε αντίθεση με τα στατιστικά μοντέλα που προδιαγράφουν τον τρόπο που εξελίσσεται η χρονοσειρά, η μηχανική μάθηση δεν απαιτεί προηγούμενες παραδοχές σχετικά με τις υποκειμενικές σχέσεις μεταξύ των μεταβλητών. Ο αλγόριθμος επεξεργάζεται τα δεδομένα και ανακαλύπτει τα πρότυπα. Οπότε πρόκειται για μια εξαγωγή γνώσης και όχι εφαρμογή προκαθορισμένων εξισώσεων. Επίσης, ένα άλλο πλεονέκτημα των μεθόδων αυτών είναι η μη γραμμικότητα οπότε καλύτερη δυνατότητα για προέκταση μη γραμμικών τάσεων και για αναγνώριση μη γραμμικών προτύπων. Ένα από τα μειονεκτήματα των μοντέλων μηχανικής μάθησης είναι ότι απαιτούν μεγάλο όγκο διαθέσιμων δεδομένων.

Η πρόβλεψη χρονοσειρών είναι ένας σημαντικός τομέας έρευνας για τη μηχανική μάθηση, ιδιαίτερα όταν η ακριβής πρόβλεψη είναι ύψιστης σημασίας, συμπεριλαμβανομένων πολλών κλάδων όπως λιανική, αλυσίδα εφοδιασμού, ενέργεια, χρηματοδότηση κ.λπ. Ενδεικτικά, στον τομέα των καταναλωτικών αγαθών, η βελτίωση της ακρίβειας της πρόβλεψης ζήτησης κατά 10-20% μπορεί να μειώσει το απόθεμα κατά 5% και να αυξήσει τα έσοδα κατά 2-3%.

Οι βασικές κατηγορίες μηχανικής μάθησης είναι η εποπτευόμενη (supervised) και η μη εποπτευόμενη μάθηση (unsupervised learning).

Εποπτευόμενη Μάθηση

Στη εποπτευόμενη μάθηση το σύστημα αξιοποιεί δεδομένα και υπάρχουσες λύσεις προκειμένου να ανακαλύψει τη σχέση που συνδέει τις μεταβλητές εισόδου (ανεξάρτητες μεταβλητές) και εξόδου (εξαρτημένες μεταβλητές). Η εποπτευόμενη μάθηση έχει σκοπό να βρει μοτίβα ανάμεσα στα δεδομένα που μπορούν να χρησιμοποιηθούν σε μια διαδικασία ανάλυσης. Η πρόβλεψη ζήτησης που είναι το αντικείμενο της παρούσας διπλωματικής εμπίπτει σε αυτή την κατηγορία εφόσον υπάρχουν προηγούμενες παρελθοντικές τιμές και ο αλγόριθμος εκπαιδεύεται ώστε να μαθαίνει από αυτά τα μοτίβα. Η συγκεκριμένη κατηγορία έχει ευρεία εφαρμογή σε μια ποικιλία επιχειρηματικών προβλημάτων, όπως προγραμματισμός εργασίας προσωπικού βάσει ζήτησης, συσχέτιση προϊόντων με προφίλ πελατείας, αναγνώριση ομιλίας, κ.λπ.

Στην επιβλεπόμενη μάθηση διακρίνονται δυο κατηγορίες προβλημάτων:

- Τα προβλήματα ταξινόμησης (classification), που σχετίζονται με την δημιουργία μοντέλων που κατηγοριοποιούν μεταξύ διακριτών κλάσεων
- Τα προβλήματα παρεμβολής (regression), που σχετίζονται με την δημιουργία μοντέλων που προβλέπουν αριθμητικές τιμές σε συνεχή διαστήματα

Υπάρχουν διάφοροι αλγόριθμοι που εντάσσονται στην κατηγορία εποπτευόμενης μάθησης, που καλύπτουν είτε προβλήματα ταξινόμησης, είτε προβλήματα παρεμβολής, είτε και τους δυο τύπους προβλημάτων. Μεταξύ άλλων είναι η γραμμική και λογιστική παλινδρόμηση, τα δέντρα απόφασης και τα απλά νευρωνικά δίκτυα.

Μη Εποπτευόμενη Μάθηση

Η μάθηση χωρίς επίβλεψη αποκαλύπτει σχέσεις που πιθανώς υπάρχουν μεταξύ των δεδομένων και τις αξιοποιεί προκειμένου να επιλύσει ένα πρόβλημα, χωρίς ωστόσο να γνωρίζει εξ αρχής κάποια λύση για αυτό. Η κατανόηση των εννοιών που κρύβονται πίσω από αυτά τα δεδομένα απαιτεί αλγόριθμους που είναι σε θέση να αρχίσουν να κατανοούν τις έννοιες αυτές, βασιζόμενοι στην ικανότητά τους να ταξινομούν τα δεδομένα βάσει των μοτίβων ή των συστάδων (clusters) που βρίσκουν. Οι αλγόριθμοι μη επιβλεπόμενης μάθησης τμηματοποιούν τα δεδομένα σε ομάδες χαρακτηριστικών (clusters). Τα προβλήματα στην μη επιβλεπόμενη μάθηση ομαδοποιούνται σε δύο κατηγορίες:

- Στα προβλήματα ομαδοποίησης (clustering): Σε ένα πρόβλημα ομαδοποίησης, ο αλγόριθμός προσπαθεί να ανακαλύψει εγγενείς ομαδοποιήσεις μεταξύ των δεδομένων στα οποία εφαρμόζεται (π.χ. ομαδοποιήσεις των πελατών με βάση την αγοραστική τους συμπεριφορά)
- Στα προβλήματα συσχέτισης (association): Σε ένα πρόβλημα συσχέτισης, ο αλγόριθμος προσπαθεί να ανακαλύψει κανόνες (rules) που χαρακτηρίζουν ένα μέρος των δεδομένων (π.χ. από ένα μεγάλο αριθμό αντικειμένων ζητείται να βρεθεί για ένα υποσύνολο τους, ποια αντικείμενα αγοράζονται μαζί)

Τα προβλήματα της μηχανικής μάθησης είναι ποικίλα και υπάρχουν πολλά είδη τεχνικών για την λύση τους. Το βασικό κριτήριο της επιλογής της κατάλληλης τεχνικής σχετίζεται με το είδος των δεδομένων που διαθέτουμε και του τελικού στόχου που θέλουμε να πετύχουμε.

Νευρωνικά δίκτυα

Ένα τεχνητό νευρωνικό δίκτυο μπορεί να αποδειχθεί χρήσιμο για μη γραμμικές διαδικασίες, όπου η συσχέτιση δεν είναι εξαρχής γνωστή και επομένως είναι δύσκολο να επιτευχθεί βέλτιστη προσαρμογή. Τα νευρωνικά δίκτυα είναι πολύ καλά στον

3.2.1 εντοπισμό προτύπων και στην ανίχνευση τάσεων στα δεδομένα υπό εξέταση.

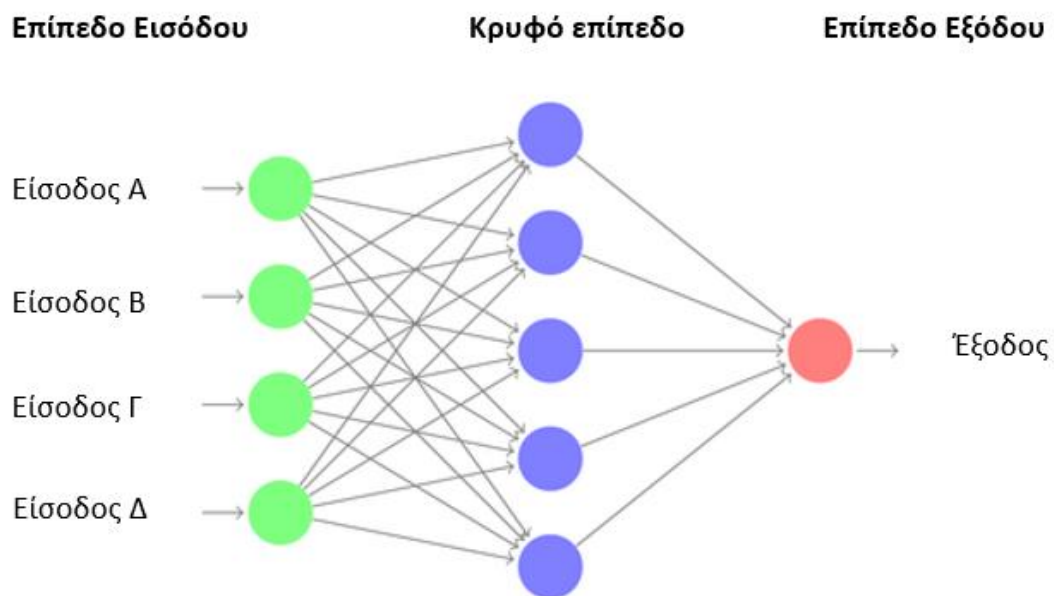
Ο αλγόριθμος εκμάθησης του τεχνητού νευρωνικού δικτύου είναι ένας αλγόριθμος μάθησης, που παίρνει την ιδέα από τη δομή των βιολογικών νευρωνικών δικτύων.

Οι υπολογισμοί βασίζονται σε ένα σύνολο εσωτερικά διασυνδεδεμένων τεχνητών νευρώνων οι οποίοι επεξεργάζονται την πληροφορία και εκτελούν υπολογισμούς επικοινωνώντας μεταξύ τους.

Χρησιμοποιούνται για τη μοντελοποίηση σύνθετων σχέσεων μεταξύ δεδομένων εισόδου και εξόδου, για την ανακάλυψη προτύπων στα δεδομένα, ή για τον εντοπισμό στατιστικής δομής σε μία άγνωστη κοινή κατανομή πιθανότητας μεταξύ των παρατηρούμενων μεταβλητών.

Η βασική ιδέα κατά την υλοποίηση του νευρωνικού είναι το φιλτράρισμα των εισόδων, δηλαδή των ανεξάρτητων μεταβλητών, μέσω ενός ή περισσότερων κρυφών επιπέδων, τα οποία αποτελούνται από κρυφούς κόμβους, πριν παραχθεί η ζητούμενη έξοδος. Συγκεκριμένα, ένα νευρωνικό δίκτυο αποτελείται από τρεις βασικές κατηγορίες νευρώνων:

- Τους νευρώνες εισόδου (input neurons), οι οποίοι δέχονται τις πληροφορίες που θα υποστούν επεξεργασία
- Τους νευρώνες εξόδου (output neurons), στους οποίους καταλήγουν τα αποτελέσματα της παραπάνω επεξεργασίας.
- Τους ενδιάμεσους νευρώνες, οι οποίοι βρίσκονται μεταξύ των νευρώνων εισόδου και εξόδου και οι οποίοι ονομάζονται κρυφοί νευρώνες (hidden neurons). Το πλήθος των κρυφών νευρώνων καθορίζονται από τον χρήστη, όπως επίσης και το πλήθος των κόμβων τους.



Εικόνα 3.3 Αρχιτεκτονική Νευρωνικών Δικτύων

Η πιο απλή και ευρύτερα χρησιμοποιούμενη κατηγορία τεχνητών νευρωνικών δικτύων είναι τα δίκτυα πρόσθιας τροφοδότησης, τα οποία ονομάζονται έτσι γιατί τα δεδομένα εισόδου τροφοδοτούνται στο δίκτυο από τους νευρώνες εισόδου και μεταφέρονται στα επόμενα στρώματα του δικτύου και στο τελευταίο στρώμα του δικτύου παράγεται η έξοδος του. Μια κατηγορία τεχνητού νευρωνικού δικτύου πρόσθιας τροφοδότησης είναι το πολυεπίπεδο δίκτυο Perceptron (MLP) το οποίο χρησιμοποιείται στην παρούσα διπλωματική και θα αναλυθεί εκτενώς παρακάτω.

3.2.1.1 Πολυεπίπεδο δίκτυο Perceptron (Multi-Layer Perceptron – MLP)

Το δίκτυο Perceptron πολλών επιπέδων είναι μια γενίκευση του απλού Perceptron και σε αντίθεση με αυτό που αποτελείται από ένα νευρώνα, το δίκτυο MLP περιλαμβάνει ένα σύνολο αισθητήρων (πηγαίοι κόμβοι) που αποτελούν το επίπεδο εισόδου, ένα ή περισσότερα κρυφά επίπεδα υπολογιστικών κόμβων και ένα επίπεδο υπολογιστικών κόμβων εξόδου. Το μοντέλο MLP είναι ένας αλγόριθμος εποπτευόμενης μάθησης δηλαδή αξιοποιεί δεδομένα και υπάρχουσες λύσεις ενός προβλήματος προκειμένου να αποκαλύψει τη σχέση που συνδέει τις μεταβλητές εισόδου (ανεξάρτητες μεταβλητές) και εξόδου (εξαρτημένες μεταβλητές). Η αρχιτεκτονική ενός δικτύου MLP εξαρτάται από διάφορες παραμέτρους, όπως ο αριθμός των νευρώνων εισόδου, ο αριθμός των κρυμμένων επιπέδων, η συνάρτηση ενεργοποίησης και ο ρυθμός μάθησης. Αυτές οι

παράμετροι παρουσιάζουν μια δυσκολία στην εύρεση της κατάλληλης αρχιτεκτονικής ενός δικτύου MLP. Στη διαδικασία της μάθησης τα βάρη που συνδέουν τους νευρώνες προσαρμόζονται ώστε να ελαχιστοποιείται η διαφορά ανάμεσα στην επιθυμητή έξοδο και στην έξοδο του δικτύου. Στην εκμάθηση του νευρωνικού, οι αλλαγές στα βάρη μπορεί να εφαρμόζονται στο δίκτυο μετά από κάθε εκμάθηση, μετά δηλαδή από κάθε πέρασμα προς τα εμπρός και προς τα πίσω. Επίσης, μπορεί να αθροίζονται οι αλλαγές από όλα τα μοτίβα εκμάθησης και το άθροισμα αυτό να εφαρμόζεται μετά από έναν πλήρη κύκλο που ονομάζεται εποχή. Ο ρυθμός μάθησης ενός δικτύου είναι μια παράμετρος που χρησιμοποιείται για να ελέγχει τον ρυθμό με τον οποίο αλλάζουν τα βάρη στο τέλος κάθε εποχής, επομένως, ανάλογα με τα δεδομένα ο ρυθμός μάθησης διαφοροποιείται προκειμένου να επιτευχθεί το βέλτιστο αποτέλεσμα. Το βασικό στοιχείο του συγκεκριμένου αλγορίθμου είναι μια υπολογιστική μονάδα η οποία λαμβάνει τον γραμμικό συνδυασμό των εισόδων της και μέσω της συνάρτησης ενεργοποίησης (activation function) παράγει την έξοδό της. Η συνάρτηση ενεργοποίησης είναι πολύ σημαντική και η επιλογή της αλλάζει ανάλογα το πρόβλημα. Η μορφή της συνήθως είναι βηματική, γραμμική, λογιστική σιγμοειδής ή υπερβολική εφαιπτομένη. Οι πιο συνηθισμένες συναρτήσεις ενεργοποίησης είναι η λογιστική σιγμοειδής και η υπερβολική εφαιπτομένη η οποία ουσιαστικά είναι μια μετατοπισμένη και κλιμακωτή εκδοχή της σιγμοειδούς.

Η εκπαίδευση ενός νευρωνικού μπορεί να γίνει με διάφορους αλγορίθμους, ο πιο δημοφιλής όμως είναι ο αλγόριθμος της προς τα πίσω διάδοσης σφάλματος (backpropagation). Ο αλγόριθμος αυτός βασίζεται στον κανόνα μάθησης διόρθωσης του λάθους. Κατά την έναρξη του αλγορίθμου τα βάρη παίρνουν τυχαίες τιμές. Έπειτα έχοντας ως στόχο την ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος ανάμεσα στις εξόδους του δικτύου και στις πραγματικές εξόδους των δεδομένων προσαρμόζονται τα νέα βάρη των συνάψεων. Σχετικά με τον αριθμό των νευρώνων εισόδου αλλά και τον αριθμό των κρυμμένων επιπέδων η βέλτιστη απόφαση εξαρτάται πάλι από τα δεδομένα προς μελέτη και θα πρέπει να γίνεται με στόχο την αποφυγή υπερπροσαρμογής ή υποπροσαρμογής, που οδηγεί σε αδυναμία του δικτύου να ανταποκριθεί σωστά σε νέα δεδομένα.

Η εκπαίδευση του τεχνητού νευρωνικού δικτύου μπορεί να γίνει ιδιαίτερα αποδοτική αν προηγηθεί επεξεργασία των δεδομένων πριν την εισαγωγή τους στο δίκτυο. Σε προβλήματα πρόβλεψης χρονοσειρών, χρησιμοποιείται η μέθοδος της κανονικοποίησης

(normalization) που έχει ως στόχο την προσαρμογή των δεδομένων σε μια κλίμακα. Για την κανονικοποίηση των δεδομένων συνήθως προτιμάται τα δεδομένα να παίρνουν τιμές μεταξύ του 0 και 1. Προκειμένου να εφαρμοστεί η διαδικασία της κανονικοποίησης χρειάζεται να είναι γνωστή η μέγιστη και η ελάχιστη τιμή της χρονοσειράς εισόδου και αντίστοιχα να οριστεί το εύρος των τιμών εισόδου max-min. Τέλος, προκειμένου να αξιολογηθούν τα αποτελέσματα που θα προκύψουν από τον συγκεκριμένο αλγόριθμο χρειάζεται να γίνει μετατροπή των δεδομένων εξόδου βάσει της μέγιστης και της ελάχιστης τιμής του δείγματος.

Δέντρα Απόφασης

3.2.2 Όπως υποδηλώνει και το όνομα, στα δέντρα απόφασης το μοντέλο που δημιουργείται έχει δενδροειδή δομή. Τα δέντρα απόφασης αποτελούν μοντέλα προβλέψεων που μπορούν να χρησιμοποιηθούν και σε προβλήματα παλινδρόμησης, όπως η πρόβλεψη ζήτησης που αναλύουμε στην παρούσα διπλωματική αλλά και σε προβλήματα ταξινόμησης. Το μοντέλο περιλαμβάνει μια σειρά από λογικές αποφάσεις, παρόμοιες με ένα διάγραμμα ροής, με κάθε κόμβο να υποδεικνύει μια απόφαση που πρέπει να ληφθεί. Τα δεδομένα που πρόκειται να ταξινομηθούν ξεκινούν από τον ριζικό κόμβο και εξετάζοντας τα γνωρίσματα που καθορίζονται από τον κόμβο, προσδιορίζεται το κλαδί που θα διασχίσουν στη συνέχεια καθώς και ο επόμενος κόμβος που θα επισκεφθούν έως ότου καταλήξουν διαδοχικά στον τελικό κόμβο ο οποίος αντιστοιχεί σε ένα φύλλο δέντρου. Κατά την παραπάνω διαδικασία, τα δεδομένα χωρίζονται με στόχο να έχουν το μικρότερο δυνατό σφάλμα. Τα δέντρα απόφασης είναι μια ευρέως χρησιμοποιούμενη τεχνική μηχανικής εκμάθησης καθώς μπορούν να εφαρμοστούν για τη μοντελοποίηση σχεδόν κάθε είδους δεδομένων και εμφανίζουν πολύ καλή απόδοση.

Στατιστικοί Δείκτες Σφάλματος

3.3 Στις προηγούμενες ενότητες έγινε η ανάλυση των διαφόρων μεθόδων πρόβλεψης και της μεθοδολογίας εξεύρεσης της καταλληλότερης μεθόδου βάσει των εκάστοτε δεδομένων. Το επόμενο στάδιο αφορά στην ανάλυση της ακρίβειας του επιλεγμένου μοντέλου πρόβλεψης. Η ακρίβεια των προβλέψεων ζήτησης στις επιχειρήσεις είναι μείζονος σημασίας, καθώς οι ανακριβείς προβλέψεις μπορεί να οδηγήσουν σε πιθανή αύξηση κόστους που σχετίζεται, για παράδειγμα, με τα αποθέματα και το προσωπικό μέχρι μείωση της ικανοποίησης των πελατών. Επιπλέον, τέτοιου είδους προβλέψεις θα μπορούσαν να έχουν ως αποτέλεσμα ο δυσαρεστημένος πελάτης να στραφεί σε ανταγωνιστικά προϊόντα. Συνεπώς, όσο πιο ακριβείς είναι οι προβλέψεις, τόσο πιο προετοιμασμένη είναι η επιχείρηση να αντιμετωπίσει την αβεβαιότητα και να πετύχει τους στόχους της. Καθώς οι αγορές αλλάζουν και ο ανταγωνισμός αυξάνεται, η ακρίβεια αυτών των προβλέψεων πωλήσεων μπορεί να προσφέρει ένα σημαντικό ανταγωνιστικό πλεονέκτημα. Πώς μπορεί όμως να μετρηθεί αυτή η ακρίβεια? Προκειμένου να υπολογίσουμε την ακρίβεια του μοντέλου, έχουν αναπτυχθεί στη βιβλιογραφία αρκετοί στατιστικοί δείκτες σφάλματος. Γενικά, το σφάλμα ορίζεται ως η διαφορά μεταξύ της πραγματικής τιμής Y και της πρόβλεψης F για μια περίοδο, όπως φαίνεται στον ακόλουθο τύπο:

$$e_i = Y_i - F_i$$

Y_i είναι η πραγματική τιμή την περίοδο i , F_i είναι η προβλεπόμενη τιμή την περίοδο i . Εξίσου σημαντική με την επιλογή κατάλληλης μεθόδου είναι και η επιλογή εκείνου του στατιστικού δείκτη σφάλματος που θα αξιολογεί την ακρίβεια της επιλεγμένης μεθόδου.

Οι στατιστικοί δείκτες σφάλματος μπορεί να υπολογισθούν είτε με τις τιμές που υπάρχουν διαθέσιμες και έχουν προβλεφθεί και από το μοντέλο (in-sample error) είτε εκ των υστέρων αφού γίνουν γνωστές οι πραγματικές τιμές οπότε είναι εφικτό να υπολογιστεί τόσο το σφάλμα του μοντέλου όσο και το πραγματικό σφάλμα (out-of-sample error). Σε γενικές γραμμές, η ακρίβεια ενός μοντέλου πρόβλεψης αξιολογείται βάσει του σφάλματος που προκύπτει επί των πραγματικών τιμών για το λόγο αυτό, κατά την ανάπτυξη ενός μοντέλου προβλέψεων συνήθως επιλέγεται ένα τμήμα των διαθέσιμων δεδομένων για εκμάθηση του μοντέλου (train set) και ένα τμήμα για την αξιολόγηση του μοντέλου (test set).

Οι βασικοί στατιστικοί δείκτες σφάλματος είναι οι ακόλουθοι:

Μέσο σφάλμα (Mean Error) - ME

Υπολογίζεται ως ο μέσος όρος των απλών σφαλμάτων. Αν ο δείκτης παίρνει θετικές τιμές, που σημαίνει ότι το Y_i (πραγματικές τιμές) είναι μεγαλύτερο του F_i (τιμές πρόβλεψης) δηλώνει απαισιοδοξία στις προβλέψεις. Το αντίθετο ισχύει στην περίπτωση που ο δείκτης είναι αρνητικός. Όσο η τιμή του δείκτη είναι κοντά στο μηδέν, τόσο τα σφάλματα είναι τυχαία και όχι συστηματικά. Το μέσο σφάλμα δείχνει την προκατάληψη (bias) της μεθόδου μας.

Ο δείκτης υπολογίζεται από τον ακόλουθο τύπο:

$$ME = \frac{1}{n} \sum_{i=1}^n (Y_i - F_i)$$

Μέσο απόλυτο σφάλμα (Mean Absolute Error) - MAE

Δηλώνει ένα μέσο μέτρο της αστοχίας της πρόβλεψης, όμως, σε αντίθεση με τον δείκτη ME, το μέσο απόλυτο σφάλμα δε δίνει έμφαση στην κατεύθυνση της πρόβλεψης. Για μια καλή πρόβλεψη, το εξαγόμενο MAE πρέπει να είναι όσο το δυνατόν μικρότερο, δηλαδή πόσο κοντά είναι η πραγματική τιμή από την προβλεπόμενη. Το μειονέκτημά του είναι πως η τιμή του εξαρτάται από την απόλυτη τιμή των δεδομένων, γεγονός που τον καθιστά αναποτελεσματικό κατά την εφαρμογή του σε σύνολα χρονοσειρών.

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - F_i|$$

Μέσο τετραγωνικό σφάλμα (Mean Squared Error) - MSE

Υπολογίζεται ως ο μέσος όρος των απλών σφαλμάτων υψωμένα στο τετράγωνο. Πρόκειται για ένα δείκτη που δίνει πολύ μεγάλο βάρος στα μεγάλα σφάλματα και αντίστοιχα μικρότερο βάρος στα μικρά σφάλματα λόγω του τετραγώνου στον υπολογισμό. Χρησιμοποιείται σε μεγάλο βαθμό σε εφαρμογές ενέργειας, για τις ώρες αιχμής, για παράδειγμα, όπου το μοντέλο χρειάζεται να έχει καλύτερη απόδοση οπότε εισάγεται το μέσο τετραγωνικό σφάλμα για να δημιουργήσει αυτή την ένταση στα μεγαλύτερα σφάλματα.

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

Ρίζα μέσου τετραγωνικού σφάλματος (Root Mean Squared Error) - RMSE

Υπολογίζεται ως η τετραγωνική ρίζα του μέσου τετραγωνικού σφάλματος επομένως έχει τις ίδιες ιδιότητες με αυτό αλλά είναι εκφρασμένο στις πραγματικές μονάδες του υπό μελέτη μεγέθους, πιο πολύ για λόγους διευκόλυνσης επικοινωνίας.

Όπως και ο προηγούμενος δείκτης, είναι ιδιαίτερα ευαίσθητος στα μεγάλα σφάλματα.

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

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

Πρόκειται για έναν από τους πιο διαδεδομένους δείκτες και βασίζεται σε ποσοστιαία σφάλματα. Χρησιμοποιείται σε περιπτώσεις όπου θέλουμε να συγκρίνουμε την ακρίβεια μιας μεθόδου πρόβλεψης που έχει εφαρμοστεί σε πολλές χρονοσειρές με διαφορετικό επίπεδο μέσης τιμής. Ένα μειονέκτημά του είναι, ότι δεν ορίζεται όταν οι πραγματικές τιμές παίρνουν μηδενικές τιμές (διακοπτόμενη ζήτηση), αφού η πραγματική ζήτηση είναι στον παρονομαστή. Λαμβάνει τιμές μεγαλύτερες ή ίσες του μηδενός, ενώ όσο μικρότερο είναι τόσο πιο ακριβής είναι η μέθοδος.

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

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

Το μειονέκτημα του δείκτη MAPE οδήγησε στη δημιουργία του αντίστοιχου δείκτη sMAPE ο οποίος έχει στον παρονομαστή το ημιάθροισμα της πραγματικής και της προβλεπόμενης τιμής οπότε αντιμετωπίζεται εν μέρει το πρόβλημα με τις μηδενικές παρατηρήσεις. Αν και το όνομα του υποδεικνύει συμμετρία, αξίζει να σημειωθεί ότι οι αισιόδοξες και οι απαισιόδοξες προβλέψεις, ίσης απόκλισης από την πραγματική τιμή, δεν καταλήγουν στην ίδια τιμή σφάλματος.

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

Μέσο απόλυτο κανονικοποιημένο σφάλμα (Mean Absolute Scaled Error) - MAsE

Πρόκειται για έναν συσχετιστικό στατιστικό δείκτη που θυμίζει το μέσο απόλυτο σφάλμα το οποίο όμως κανονικοποιείται με τη μέση τιμή των διαφορών πρώτου βαθμού της χρονοσειράς. Ο παρονομαστής ουσιαστικά αποτελεί το μέσο απόλυτο σφάλμα της μεθόδου Naïve, η οποία σαν πρόβλεψη για την επόμενη περίοδο, δίνει την πραγματική τιμή της τρέχουσας περιόδου.

$$MAsE = \frac{\frac{1}{n} \sum_{i=1}^n |Y_i - F_i|}{\frac{1}{n-1} \sum_{i=2}^n |Y_i - Y_{i-1}|}$$

Οι στατιστικοί δείκτες σφαλμάτων θα μπορούσαν να διαχωριστούν στις ακόλουθες κατηγορίες

- Δείκτες που εξαρτώνται από την κλίμακα των δεδομένων
- Δείκτες που βασίζονται σε ποσοστιαία σφάλματα
- Δείκτες που βασίζονται σε σχετικά σφάλματα
- Σχετικοί δείκτες
- Κανονικοποιημένοι δείκτες

Συστήματα διαχείρισης επιχειρησιακών πόρων (Enterprise Resource Planning - ERP)

3.4 Το ακρωνύμιο ERP προέρχεται από τα αρχικά των λέξεων Enterprise Resource Planning, το οποίο μπορεί να αποδοθεί στα ελληνικά ως «σύστημα διαχείρισης επιχειρησιακών πόρων». Τα συστήματα ERP είναι ολοκληρωμένα συστήματα πληροφορικής, τα οποία καλύπτουν όλες τις λειτουργικές περιοχές μιας επιχείρησης, συμπεριλαμβανομένου του σχεδιασμού και ανάπτυξης προϊόντων, της παραγωγής, των χρηματοοικονομικών, τη διαχείριση ανθρώπινου δυναμικού, των πωλήσεων και του μάρκετινγκ, ώστε ενοποιώντας όλες τις διαδικασίες της επιχείρησης να ικανοποιηθούν οι στόχοι της. Πρωτοεμφανίστηκε τη δεκαετία του '60, από την εταιρεία IBM επιφορτισμένο με τη διαχείριση πρώτων υλών (MRP) για τις ανάγκες του κλάδου της βιομηχανίας. Με τα χρόνια, σημειώνει ραγδαία πρόοδο εντάσσοντας όλες τις νέες τεχνολογίες που αναδύονται ραγδαία για να καλύψουν τις ανάγκες των εταιρειών. Μέσω της δια λειτουργικότητας που προσφέρει, επιτυγχάνεται η ταχύτατη, ακριβής και έγκαιρη μετάδοση της πληροφορίας στο εσωτερικό της επιχείρησης. Αυτή η πληροφορία μπορεί να αφορά σε κόστος, κέρδη, υλικά και διάφορες άλλες επιχειρηματικές διαδικασίες. Μέσω ενός ERP, τα στελέχη διοίκησης έχουν τη δυνατότητα να βλέπουν σε πραγματικό χρόνο μια ολιστική εικόνα της επιχείρησης για να λάβουν αποφάσεις που αφορούν μελλοντικές κινήσεις, καθορίζοντας τη στρατηγική και τη θέση της επιχείρησης στην αγορά.

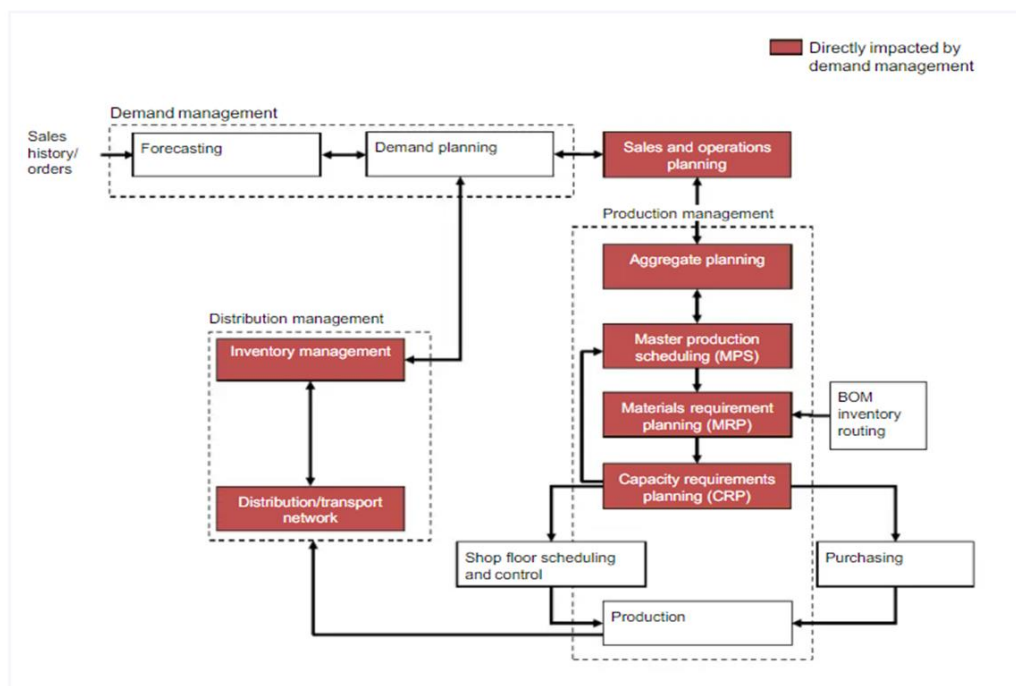


Εικόνα 3.4 Σύστημα Διαχείρισης Επιχειρησιακών Πόρων (ERP)

Τα συστήματα ERP είναι δομημένα σε λειτουργικά υποσυστήματα (functional modules). Τα κύρια υποσυστήματα παρουσιάζονται παρακάτω, ωστόσο, δεν είναι απαραίτητο κάθε εταιρική εφαρμογή συστήματος να τα περιλαμβάνει όλα.

1. Οικονομικής διαχείρισης
2. Διαχείρισης της αλυσίδας προμηθειών
3. Προγραμματισμού Παραγωγής
4. Μάρκετινγκ και πωλήσεων
5. Διαχείρισης αποθεμάτων
6. Διανομών
7. Διαχείρισης ανθρώπινων πόρων
8. Διαχείρισης των σχέσεων με πελάτες (CRM)

Στο παρακάτω διάγραμμα φαίνεται πώς επηρεάζεται η πλειοψηφία των υποσυστημάτων από τη διαδικασία πρόβλεψης. Η παρούσα διπλωματική θα επικεντρωθεί στο υποσύστημα του προγραμματισμού παραγωγής και εν προκειμένω στη διαδικασία προγραμματισμού απαιτήσεων υλικών(MRP) στην οποία τρέχουν τα μοντέλα πρόβλεψης ζήτησης προτείνοντας την ποσότητα παραγωγής για τον κάθε κωδικό.



Εικόνα 3.5 Αντίκτυπο της πρόβλεψης στις λειτουργικές περιοχές της επιχείρησης

Υποσύστημα προγραμματισμού παραγωγής

Το υποσύστημα παραγωγής ανταλλάσσει πληροφορίες με το υποσύστημα οικονομικής διαχείρισης, το υποσύστημα πωλήσεων και μάρκετινγκ, προμηθειών καθώς και εφοδιαστικής αλυσίδας. Καλύπτει λειτουργίες οι οποίες σχετίζονται με την

3.4.1 παραγωγική διαδικασία. Κάποιες από αυτές είναι οι παρακάτω:

- Προγραμματισμός παραγωγής για μηχανήματα με την βέλτιστη αξιοποίηση όλων των διαθέσιμων πόρων.
- Σχεδιασμός με βάση τις προβλέψεις πωλήσεων.
- Προγραμματισμός υλικών (MRP) με βάση τη χωρητικότητα και τη διαθεσιμότητα του μηχανήματος. Αποδοτικότητα μηχανής, διαθεσιμότητα πρώτων υλών, χρόνος παράδοσης.
- Δυνατότητα ανάκλησης του σχεδίου παραγωγής για αλλαγή παραμέτρων εισόδου, προτεραιότητα παραγωγής ή ποσότητα.
- Ανάλυση της αποδοτικότητας και της χρήσης της μηχανής.
- Αυτόματη παραγωγή βασικού χρονοδιαγράμματος παραγωγής(MRS) και αιτήσεις αγοράς για την ολοκλήρωση του προγράμματος.
- Δημιουργία αιτήματος για διαδικασίες που πρέπει να ανατεθούν σε υπεργολαβία.
- Κράτηση ποσότητας για παραγωγή.
- Έλεγχος και κοστολόγηση παραγωγής
- Αυτόματη δημιουργία παραγγελιών για παραγωγή.
- Δημιουργία αναφορών που σχετίζονται με την παραγωγή

Ένα βασικό στοιχείο του υποσυστήματος παραγωγής είναι το βασικό αρχείο υλικών, το οποίο αποτελείται από τις παρακάτω δομές:

- Πίνακα υλικών (BOM). Ο πίνακας υλικών είναι μια λίστα των συστατικών που απαρτίζουν ένα προϊόν ή μια συναρμολόγηση μέχρι οποιοδήποτε αριθμό επιπέδων.
- Στη λίστα αυτή περιέχεται ο αριθμός αναλυτικών γραμμών κάθε συστατικού, η ποσότητα και η μονάδα μέτρησης.
- Κέντρο εργασίας. Κέντρο εργασίας είναι ο χώρος εκτέλεσης μιας λειτουργίας ή μιας δραστηριότητας μέσα σε μια εγκατάσταση.

- Φασεολόγιο. Περιγράφει τις διάφορες φάσεις μιας παραγωγικής διαδικασίας.
- Κέντρο κόστους

Λύσεις λογισμικού ERP

SAP SE (S/4HANA, SAP R3, SAP Business One)

Η εταιρεία ιδρύθηκε τον Απρίλιο 1972 στην πόλη Mannheim της Γερμανίας με την επωνυμία System Analyse und Programmentwicklung από πέντε συμβούλους και προγραμματιστές που παλιότερα δούλευαν στην IBM της Γερμανίας.

Σκοπός της ίδρυσης της εταιρείας ήταν η ανάπτυξη ενός πρότυπου λογισμικού εφαρμογών που επεξεργάζεται δεδομένα σε πραγματικό χρόνο. Η πραγματοποίηση αυτού του στόχου έκανε τη SAP τον μεγαλύτερο προμηθευτή επιχειρησιακού λογισμικού παγκοσμίως, και οι ERP λύσεις της εταιρίας απευθύνονται τόσο σε μεγάλες επιχειρήσεις (S/4HANA) όσο και σε μεσαίες επιχειρήσεις (SAP Business One) και είναι οι πιο γνωστές και διαδεδομένες παγκοσμίως. Η παρουσία της SAP στην Ελλάδα άλλαξε την έννοια του επιχειρησιακού λογισμικού στην ελληνική αγορά με την απόλυτη προσαρμογή του συστήματος SAP στα δεδομένα και στις απαιτήσεις της ελληνικής πραγματικότητας (ελληνικοποίηση).

Η στροφή των εταιριών κατασκευής λογισμικού τα τελευταία χρόνια στο cloud computing έχει αναγκάσει όλους τους μεγάλους προμηθευτές λογισμικού να προσαρμόσουν τις λύσεις τους στην νέα πραγματικότητα. Η SAP ανέπτυξε τη σουίτα εφαρμογών S/4HANA, που αποτελεί ουσιαστικά την εξέλιξη του SAP R3 και αποτελεί ένα έξυπνο και ολοκληρωμένο σύστημα ERP που χρησιμοποιεί το SAP HANA, δηλαδή την δική του in-memory βάση δεδομένων. Η βάση δεδομένων εντός μνήμης επιτυγχάνει γρήγορη επεξεργασία δεδομένων σε πραγματικό χρόνο. Επίσης, το λογισμικό προσφέρει τη δυνατότητα για δημόσιο/ιδιωτικό cloud ή υβριδική ανάπτυξη και έχει ενσωματωμένες λύσεις AI, αναλύσεις και ευφυή αυτοματισμός διεργασιών.

Oracle Corporation (Oracle Cloud ERP, Oracle Business Intelligence)

Στην ίδια κατηγορία προμηθευτών ανήκει και η Oracle, η αμερικάνικη πολυεθνική που ιδρύθηκε το 1977 με το όνομα Software Development Laboratories (SDL) από τους Larry Ellison, Bob Miner και Ed Oates. Η εταιρία μετονομάστηκε σε Oracle Systems Corporation το 1982, για να συμβαδίσει περισσότερο με το βασικό προϊόν της, Oracle Database (Oracle RDBMS), λογισμικό διαχείρισης σχεσιακών βάσεων δεδομένων

(database management system). Αν και η εταιρία δεν έχει ως κύριο προϊόν της τις εφαρμογές ERP, έχει καταφέρει να κατέχει το δεύτερο μεγαλύτερο μερίδιο αγοράς σε αυτή την κατηγορία παγκοσμίως, και να είναι ο κύριος ανταγωνιστής της SAP εδώ και πάνω από μία δεκαετία. Η νέα σουίτα εφαρμογών της Oracle ERP Cloud απευθύνεται σε όλες τις επιχειρήσεις και προσαρμόζεται στο μέγεθος και τις ανάγκες τους και όπως και τα υπόλοιπα σύγχρονα ERP συστήματα, προσφέρει κάθε λύση σε μορφή εφαρμογών που η καθεμία εξυπηρετεί συγκεκριμένη λειτουργία και τμήμα του οργανισμού αλλά ενσωματώνονται απόλυτα μεταξύ τους ώστε να προσφέρουν στο χρήστη μια απρόσκοπτη και ενιαία εμπειρία χρήσης. Επίσης, μια άλλη λύση της Oracle που απαντάει στην ολοένα και αυξανόμενη τάση των επιχειρήσεων για εκμετάλλευση των δεδομένων είναι το Oracle Business Intelligence. Αποτελεί μια ολοκληρωμένη BI λύση η οποία περιλαμβάνει προηγμένη τεχνολογία data warehousing καθώς Dashboards και ad hoc reporting και online analysis.

Microsoft Corporation (Dynamics 365 Suite)

Ίσως ο πιο γνωστός προμηθευτής λογισμικού και προϊόντων πληροφορικής παγκοσμίως, η Microsoft ιδρύθηκε το 1975 από τους Bill Gates και Paul Allen. Το 1980 η εταιρία μπήκε στην αγορά των λειτουργικών συστημάτων, και καθιερώθηκε με το MS-DOS που την έφερε στην πρώτη θέση, ξεπερνώντας όλους τους ανταγωνιστές της. Η επανάσταση ήρθε το 1985, με την κυκλοφορία των Microsoft Windows, του πρώτου λειτουργικού συστήματος με γραφικό περιβάλλον ενώ άλλο ένα προϊόν που έφερε την Microsoft σχεδόν μέσα σε κάθε σπίτι και κάθε επιχείρηση είναι το Microsoft Office, μια σουίτα ανεξάρτητων εφαρμογών για επεξεργασία κειμένου και άλλων δεδομένων, που κυκλοφόρησε για πρώτη φορά το 1990.

Αν και η Microsoft έχει επενδύσει σε πολλούς τομείς της πληροφορικής από την ίδρυσή της και μετά, τόσο σε λογισμικό όσο και σε εξοπλισμό (υπολογιστές, περιφερειακά υπολογιστών, κονσόλες βιντεοπαιχνιδιών, κλπ.), στην αγορά των ERP μπήκε μόλις το 2002, όταν εξαγόρασε την εταιρία Navision A/S, αποτέλεσμα της συγχώνευσης των εταιριών Damgaard Data και Navision Software A/S, που είχαν αναπτύξει το Axapta και Navision αντίστοιχα. Το 2011 η Microsoft κυκλοφόρησε το Dynamics AX 2012, ένα ERP που απευθυνόταν σε μεγάλες και πολύ μεγάλες επιχειρήσεις και οργανισμούς. Το 2016 η εταιρία έβγαλε την νέα έκδοση του προϊόντος, επανασχεδιάζοντάς το ουσιαστικά ώστε να είναι πλήρως λειτουργικό ως cloud εφαρμογή με νέο UI. Η νέα

έκδοση μετονομάστηκε σύντομα σε Dynamics 365 for Finance & Operations και αποτέλεσε τη ναυαρχίδα της εταιρίας στο επιχειρησιακό λογισμικό, δημιουργώντας μία σειρά από cloud εφαρμογές, ανεξάρτητες αλλά πλήρως διασυνδεδεμένες και με κοινό περιβάλλον χρήστη. Πλέον το Dynamics 365 έχει μπει δυναμικά στο παιχνίδι της αγοράς ERP, κατακτώντας την έκτη θέση στην παγκόσμια κατάταξη, ενώ το 80% των νέων πελατών επιλέγει την υλοποίηση του ERP στο cloud.

Το Microsoft Dynamics 365 έχει καταφέρει να εξαλείψει την πολυπλοκότητα των διαφορετικών συστημάτων CRM και ERP, δημιουργώντας σύγχρονες, modular επιχειρηματικές εφαρμογές που λειτουργούν μαζί σε μια ενιαία πλατφόρμα, παρέχοντας στις επιχειρήσεις την ευελιξία να υιοθετούν την κατάλληλη τεχνολογία όταν την χρειάζονται, βελτιώνοντας την παραγωγικότητά τους.

Entersoft A.E. (Entersoft Business Suite)

Το Entersoft Business Suite είναι η ολοκληρωμένη και ενοποιημένη σουίτα επιχειρηματικού λογισμικού της Entersoft για ERP, CRM, Retail, Mobile, E-Commerce και Business Intelligence εφαρμογές. Παρέχει τη δυνατότητα διαχείρισης και παρακολούθησης πολλαπλών εταιρικών διαστάσεων και υποστηρίζει πολυεταιρική οργάνωση σε ενιαία Βάση Δεδομένων. Η εταιρεία, η οποία είναι αμιγώς ελληνική, ιδρύθηκε το 2002 από στελέχη που ήταν βασικοί μέτοχοι και ιδρυτές της εταιρείας Computer Logic και προσφέρει μια ολοκληρωμένη σουίτα εφαρμογών για μεσαίες και μεγάλες επιχειρήσεις. Σταδιακά, η εταιρία εισέρχεται και στις αγορές περισσότερων επιχειρησιακών εφαρμογών και μέχρι σήμερα έχει εδραιωθεί στην Ελλάδα ως ένας από τους κυριότερους παίκτες της αγοράς πληροφορικής.

3.4.3

Προγραμματισμός Απαιτήσεων Υλικών (Materials Requirements Planning - MRP)

Ο Προγραμματισμός Απαιτήσεων Υλικών - MRP αποτελεί μέρος της διαδικασίας προγραμματισμού παραγωγής. Η ανάγκη των εταιρειών, τη δεκαετία του 1960, να μηχανογραφήσουν τις πολύπλοκες λειτουργίες τους και συγκεκριμένα την ικανοποίηση του στόχου που είχαν θέσει εκείνη τη χρονική περίοδο να διατηρούν μεγάλο αριθμό αποθεμάτων προκειμένου να είναι σε θέση να ικανοποιούν οποιαδήποτε ζήτηση από τους πελάτες και ταυτόχρονα να είναι ανταγωνιστικοί, είχε σαν αποτέλεσμα τη δημιουργία εφαρμογών μηχανογράφησης με στόχο την

αποτελεσματική οργάνωση και διαχείριση των αποθεμάτων και της αποθήκης των επιχειρήσεων. Στα τέλη της δεκαετίας του 1960 και στις αρχές της επόμενης δεκαετίας, εμφανίστηκαν τα συστήματα σχεδιασμού απαιτήσεων υλικών (Material Requirements Planning - MRP) τα οποία αποτέλεσαν την αφετηρία όλων των εξελίξεων. Στόχος των MRPs ήταν η υλοποίηση μιας ολοκληρωμένης λύσης στο επιχειρηματικό περιβάλλον. Η ανάγκη που αποτέλεσε την αφετηρία για την ανάπτυξη του MRP ήταν η αδυναμία των εταιρειών να 4 διατηρούν μεγάλα αποθέματα. Έτσι χρησιμοποιώντας έναν υπολογιστή και λαμβάνοντας υπόψη τις πληροφορίες για τις απαιτήσεις – ζητήσεις των προϊόντων καθώς και τα υλικά που χρειάζονται για τα υπό κατασκευή προϊόντα, μπορούσε να γίνει προγραμματισμός των απαιτήσεων των υλικών. Κατά την διάρκεια της δεκαετίας του 1980 εμφανίστηκε το MRP-II (Manufacturing Resources Planning), ένα σύστημα MRP που συνέδεσε μεταξύ τους τα κυκλώματα προγραμματισμού παραγωγής, του ελέγχου παραγωγής, της κοστολόγησης, και των προμηθειών. Την ίδια δεκαετία η κατακόρυφη αύξηση της υπολογιστικής ισχύος ενισχύει την ερευνητική προσπάθεια για επιχειρηματική ολοκλήρωση (enterprise integration), η οποία χρησιμοποιεί ως τεχνολογικό υπόβαθρο τις βάσεις δεδομένων (databases) και προσπαθεί να ενώσει τις βασικές επιχειρηματικές διαδικασίες και κυρίως την οικονομική διαχείριση και την παραγωγή. Πλέον, ένα ολοκληρωμένο MRP λογισμικό εποπτεύει και διαχειρίζεται ολόκληρο το σύστημα, από την παραγγελία των υλικών μέχρι τον χρονικό προγραμματισμό του συστήματος, τον έλεγχο αποθέματος, τα οικονομικά και λογιστικά ζητήματα.

3.4.4

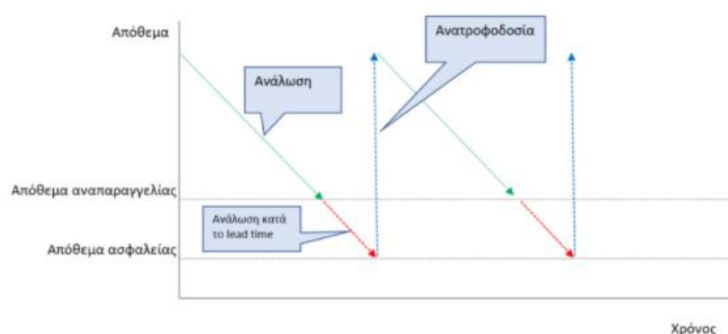
Πως λειτουργεί το MRP

Το MRP δημιουργεί προτάσεις αναπλήρωσης για όλα τα υλικά που συμμετέχουν στην παραγωγική διαδικασία (παραγόμενα, ενδιάμεσα, συμπαράγωγα, Α και Β ύλες, υλικά συσκευασίας). Μέσω του MRP ορίζονται ημερομηνίες διαθεσιμότητας των εντολών παραγωγής και των αιτήσεων αγοράς λαμβάνοντας υπόψη το συνολικό χρόνο διέλευσης των υλικών (lead time), ο οποίος εξαρτάται από χρόνους παραγωγής και αναμονής, κανόνες σχηματισμού μερίδας παραγωγής (lot size) και άλλα κριτήρια. Η εκτέλεση του προγραμματισμού μπορεί να γίνει τόσο συνολικά, για όλες τις εγκαταστάσεις και υλικά, όσο και επιλεκτικά για συγκεκριμένες εγκαταστάσεις και ομάδες υλικών. Με τη βοήθεια μεθόδων και εργαλείων που παρέχονται από το σύστημα είναι επίσης δυνατή η αξιολόγηση και εξισορρόπηση της δυναμικότητας

(capacity evaluation & leveling). Η ζήτηση των τελικών προϊόντων καθορίζει και τις ανάγκες σε πρώτες ύλες και υλικά συσκευασίας. Η σχέση μεταξύ τελικών προϊόντων και απαιτούμενων υλικών (ή ενδιάμεσων) καθορίζεται από την Συνταγή ή BOM (Bill of Materials). Ένα αυτοματοποιημένο MRP λαμβάνει υπόψιν του τις παραμέτρους και προτείνει αγορές υλικών, τέτοιες ώστε να καλύπτονται οι ανάγκες και οι πιθανές αποκλίσεις.

Παράμετροι του MRP

- Πρόβλεψη πωλήσεων-ζήτηση, η οποία μεταφράζεται σε αναλώσεις υλικών
- Ελάχιστη ποσότητα αναπαραγωγής (batch size)
- Ημερομηνία λήξης (αν υπάρχει) ή υποβάθμισης του υλικού
- Χρόνος απόκρισης του προμηθευτή (lead time) μαζί με χρόνο μεταφοράς των υλικών.
- Απόθεμα ασφαλείας



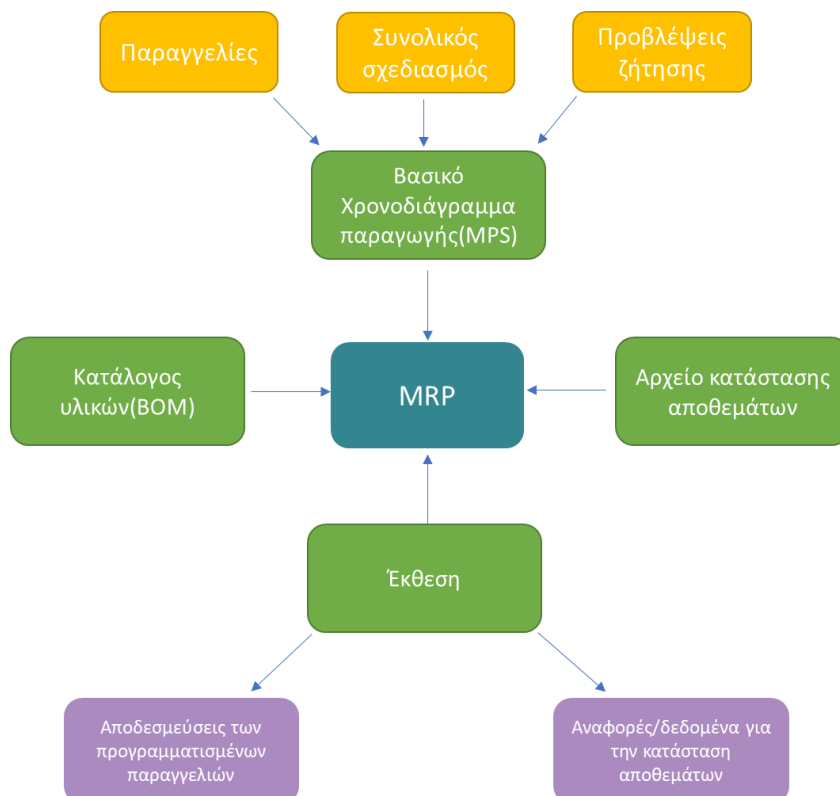
Εικόνα 3.6 Τρόπος λειτουργίας του MRP

Προκειμένου να διαμορφωθούν τα διάφορα εναλλακτικά, συμπληρωματικά ή διαδοχικά πλάνα προγραμματισμού, υπάρχει μία σειρά από εργασίες προετοιμασίας που πρέπει να προηγηθούν, όπως: ο ορισμός των παραγωγικών μονάδων, των παραγωγικών πόρων, των εργάσιμων ημερών των παραγωγικών μονάδων της εταιρείας κ.α. Το Βασικό Χρονοδιάγραμμα Παραγωγής (Master Production Schedule – MPS) στηρίζεται στην αναγνώριση μελλοντικών απαιτήσεων σε τελικά προϊόντα βάσει:

- Παραγγελιών υπαρκτών πελατών με βάση την επιθυμητή-συμφωνημένη ημερομηνία παράδοσης
- Πρόβλεψης ζήτησης
- Έγκαιρης δημιουργίας αποθέματος

Όπως και στον προγραμματισμό παραγωγής, έτσι και στον προγραμματισμό αναγκών αποθεμάτων υπάρχουν κάποια κριτήρια από όπου επιλέγει ο χρήστης τις απαιτήσεις αναπλήρωσης αποθεμάτων.

- ο Εγγραφές προγραμματισμού
- ο Προϋπολογισμός πωλήσεων
- ο Παραγγελίες πώλησης



Εικόνα 3.7 Διαδικασία προγραμματισμού της παραγωγής

Ενσωματωμένα μοντέλα πρόβλεψης σε ERP συστήματα

Κάθε σύστημα ERP έχει ενσωματωμένο τα δικά του μοντέλα πρόβλεψης. Σε αυτή τη παράγραφο θα παρουσιάσουμε, ενδεικτικά, τα μοντέλα που χρησιμοποιούν δυο κορυφαία συστήματα ERP της εταιρείας SAP και της εταιρείας Microsoft.

3.4.5 Microsoft Dynamics 365

- Κινητός Μέσος Όρος
- ETS - Εκθετική Εξομάλυνση
- Μέθοδος Croston
- ARIMA
- Πολλαπλή Γραμμική Παλινδρόμηση
- STL – Μέθοδος Αποσύνθεσης
- Συνδυασμός μεθόδων

SAP ERP

- Κινητός Μέσος Όρος
- Σταθμισμένος Κινητός Μέσος Όρος
- Εκθετική Εξομάλυνση - Μοντέλο Σταθερού Επιπέδου
- Εκθετική Εξομάλυνση - Μοντέλο Γραμμικής Τάσης
- Εποχιακή Εξομάλυνση
- Γραμμική Παλινδρόμηση

Πειραματική Διαδικασία

Καταγραφή του προβλήματος

4 Η συνεχιζόμενη μεταβλητότητα της ζήτησης σε συνδυασμό με τη δυναμική της αγοράς
4.1 αναγκάζει τις εταιρείες να αναπτύξουν πιο ολοκληρωμένες διαδικασίες διαχείρισης
ζήτησης, οι οποίες απαιτούν εξελιγμένες τεχνολογίες όπως τη χρήση Analytics και IoT
για να επιτύχουν τους στόχους τους αυξάνοντας τον τζίρο. Αυτές οι αλλαγές στη
δυναμική της αγοράς έχουν φέρει στο επίκεντρο τη διαδικασία πρόβλεψης και την
ενορχήστρωση της καλύτερης απόκρισης της ζήτησης. Επιπλέον, οι συρρικνωμένοι
κύκλοι ζωής των προϊόντων, σε συνδυασμό με μια αυξανόμενη ανταγωνιστική πίεση,
αυξάνουν το αντίκτυπο ενός λάθους στην πρόβλεψη ζήτησης.

Η διαμόρφωση της ζήτησης χρησιμοποιεί παράγοντες, όπως η τιμή, η κυκλοφορία
νέων προϊόντων, οι προωθήσεις εμπορίου και πωλήσεων, η διαφήμιση και τα
προγράμματα μάρκετινγκ, εκτός από άλλες σχετικές πληροφορίες πωλήσεων και
μάρκετινγκ, για να επηρεάσουν τι και πόσο θα αγοράσουν οι καταναλωτές.
Συνεπώς, ο προγραμματισμός παραγωγής πρέπει να είναι σχεδιασμένος με τέτοιο
τρόπο ώστε να ανταποκρίνεται άμεσα σε αυτές τις αλλαγές.
Μια επιχείρηση που μένει πίσω στο στίβο του προγραμματισμού παραγωγής,
διακινδυνεύει να αποτύχει να παράξει τα απαιτούμενα προϊόντα στην ποσότητα και
στον χρόνο που απαιτείται. Ο πελάτης ενδιαφέρεται να παραλάβει το προϊόν που έχει
παραγγείλει στον συμφωνημένο χρόνο και με την συμφωνημένη διαδικασία. Ένα
ολοκληρωμένο σύστημα προγραμματισμού παραγωγής αποφεύγει τη δέσμευση
σημαντικών κεφαλαίων σε αχρησιμοποίητες πρώτες ύλες και υλικά συσκευασίας, ή και
έτοιμα προϊόντα. Ένα σύγχρονο και αυτοματοποιημένο μοντέλο πρόβλεψης ζήτησης
πρέπει να είναι σε θέση να ανταποκρίνεται σε αυτές τις απαιτήσεις αναλύοντας με τον
κατάλληλο τρόπο τα δεδομένα έτσι ώστε να αυξάνεται η αποδοτικότητα της εταιρείας.
Στόχος της παρούσας διπλωματικής είναι η βελτιστοποίηση της πρόβλεψης ζήτησης
προκειμένου να επιτευχθεί πιο στοχευμένος προγραμματισμός παραγωγής για την
 κάλυψη των αναγκών μιας μεσαίου μεγέθους εταιρείας τροφίμων που παράγει
συσκευασμένα καταναλωτικά προϊόντα (CPG).

To R Studio ως εργαλείο προβλέψεων

Η επεξεργασία των δεδομένων και η ανάπτυξη των μοντέλων της παρούσας διπλωματικής εργασίας, έγινε σε γλώσσα προγραμματισμού R μέσω του περιβάλλοντος Rstudio.

4.2

Η R είναι ένα λογισμικό ανοιχτού κώδικα (open source) και αποτελεί ταυτόχρονα μια ισχυρή αντικειμενοστραφή γλώσσα προγραμματισμού που προτιμάται για την ανάλυση και τον υπολογισμό διαφόρων στατιστικών μεθόδων, όπως τα γραμμικά και μη γραμμικά μοντέλα, τις χρονοσειρές ενώ τελευταία εφαρμόζεται και για αλγορίθμους μηχανικής μάθησης. Σχεδιάστηκε ως μια στατιστική πλατφόρμα για αποτελεσματικό χειρισμό δεδομένων όπως καθαρισμός, ανάλυση και αναπαράσταση δεδομένων διαθέτοντας μια μεγάλη γκάμα τεχνικών δημιουργίας γραφημάτων.

Η R δημιουργήθηκε τη δεκαετία του 1990 από τους Ross Ihaka και Robert Gentleman και η υλοποίηση της βασίστηκε στη γλώσσα προγραμματισμού S. Στην αρχή, δεν συγκαταλεγόταν στις δημοφιλείς γλώσσες, πλέον όμως, οι δυνατότητές της έχουν επεκταθεί αρκετά και έχει τεράστια απήχηση στην επιστήμη των δεδομένων. Χρησιμοποιείται ευρέως ως εκπαιδευτική γλώσσα και ως ερευνητικό εργαλείο και μπορεί να μεταγλωττιστεί και να εκτελεστεί σε μεγάλη γκάμα λειτουργικών συστημάτων, όπως τα Windows, Unix και MacOS. Όλο και περισσότερος κόσμος που ασχολείται με την επιστήμη δεδομένων μεταβαίνει στο R από συμβατικά στατιστικά πακέτα όπως SPSS, SAS και Stata, αφενός γιατί είναι δωρεάν διαθέσιμη και αφετέρου λόγω της ευελιξίας και των δυνατοτήτων οπτικοποίησης δεδομένων που προσφέρει.

Αναφορικά με το περιβάλλον RStudio, είναι ένα ελεύθερο προγραμματιστικό περιβάλλον για την γλώσσα R, το οποίο παρέχει μια οργανωμένη διάταξη και διάφορες πρόσθετες επιλογές. Το πρόγραμμα αυτό παρέχει μια ευρεία γκάμα δυνατοτήτων για την δημιουργία, ανάλυση, πρόβλεψη και απεικόνιση χρονοσειρών.

Οι χρήστες της γλώσσας R την προτιμούν, ενδεικτικά, για τους παρακάτω λόγους.

Δωρεάν και ανοιχτού κώδικα

Η γλώσσα προγραμματισμού R είναι ανοιχτού κώδικα και δεν περιορίζεται αυστηρά στα λειτουργικά συστήματα. Όντας ανοιχτής πηγής, η γλώσσα R καλύπτεται από το GNU General Public License Agreement γεγονός που την καθιστά εξαιρετικά οικονομική για ένα έργο οποιουδήποτε μεγέθους. Δεδομένου ότι είναι ανοιχτού κώδικα, οι εξελίξεις στο R συμβαίνουν σε ταχεία κλίμακα και η κοινότητα των

προγραμματιστών είναι τεράστια. Ο ερευνητής ή οποιοσδήποτε προγραμματιστής R μπορεί να εξετάσει τον πηγαίο κώδικα και να διορθώσει σφάλματα και δεν υπάρχει η ανάγκη αναμονής μιας πιο ενημερωμένης έκδοσης. Ενώ η R παρέχει μια πληθώρα δυνατοτήτων για τη στατιστική μοντελοποίηση, ωστόσο, οι προγραμματιστές μπορούν εύκολα να γράψουν το δικό τους λογισμικό και να το διανείμουν με τη μορφή πρόσθετων πακέτων (packages). Λόγω της σχετικής ευκολίας δημιουργίας αυτών των πακέτων, έχει δημιουργηθεί μια τεράστια συλλογή βιβλιοθηκών-πακέτων. Επίσης, η γλώσσα R παρέχει και μια εκτεταμένη λίστα πακέτων μηχανικής μάθησης, κάνοντας αυτόν τον κλάδο πιο εύκολο και προσιτό. Ένα παράδειγμα, που θα χρησιμοποιηθεί και στην παρούσα διπλωματική είναι το πακέτο *Random Forest* για τη δημιουργία δέντρων αποφάσεων.

Επικοινωνία με άλλες γλώσσες

Καθώς όλο και περισσότεροι χρήστες στράφηκαν στη γλώσσα R για τις αναλύσεις τους, άρχισαν να προσπαθούν να συνδυάσουν το R με τις προηγούμενες ροές εργασίας τους, γεγονός που οδήγησε σε ένα σύνολο πακέτων για τη σύνδεση του R με συστήματα αρχείων, βάσεις δεδομένων και άλλες εφαρμογές. Πολλά από αυτά τα πακέτα έχουν πλέον ενσωματωθεί στη βασική εγκατάσταση του R. Αρχικά υπήρχε δυνατότητα σύνδεσης μόνο με τις γλώσσες Fortran και C, καθώς όμως η κοινότητα μεγάλωνε, και άλλες δημοφιλείς γλώσσες όπως C++, Java και Python συνδέθηκαν με την R. Επίσης, καθώς πολλοί στατιστικολόγοι χρησιμοποιούσαν και άλλα εμπορικά προγράμματα (SAS, SPSS), δημιουργήθηκαν εργαλεία για να διαβάζουν δεδομένα από αυτά τα προγράμματα.

Εξειδίκευση στη στατιστική ανάλυση και ευκολία στη χρήση

Πρόκειται για μια γλώσσα σχεδιασμένη ειδικά για στατιστική ανάλυση, συνεπώς, όλες οι βιβλιοθήκες της R επικεντρώνονται στη δημιουργία μιας πιο εύκολης, προσιτής και λεπτομερούς ανάλυσης δεδομένων. Συγκεκριμένα, διατίθενται πολλά πακέτα που απλοποιούν τη διαδικασία προετοιμασίας των δεδομένων. Η διαδικασία καθαρισμού και μετατροπής δεδομένων είναι επίσης εξαιρετικά απλή και μπορεί να δημιουργηθεί ένα καθαρό dataset χωρίς κενές τιμές με μια γραμμή κώδικα. Επιπρόσθετα, παρέχεται μια πληθώρα πακέτων για προηγμένες απεικονίσεις των δεδομένων

Περιγραφή δεδομένων

Πριν προχωρήσουμε στην ανάλυση της μεθοδολογίας που ακολουθήθηκε θα πρέπει να αναφερθούμε στο είδος των δεδομένων που χρησιμοποιήθηκαν στα πλαίσια της παρούσας διπλωματικής. Η εν λόγω διπλωματική, βασίστηκε στα μηνιαία δεδομένα 4.3 πωλήσεων της εταιρείας τροφίμων, για την περίοδο από Ιανουάριο του 2017 έως και Δεκέμβριο του 2019.

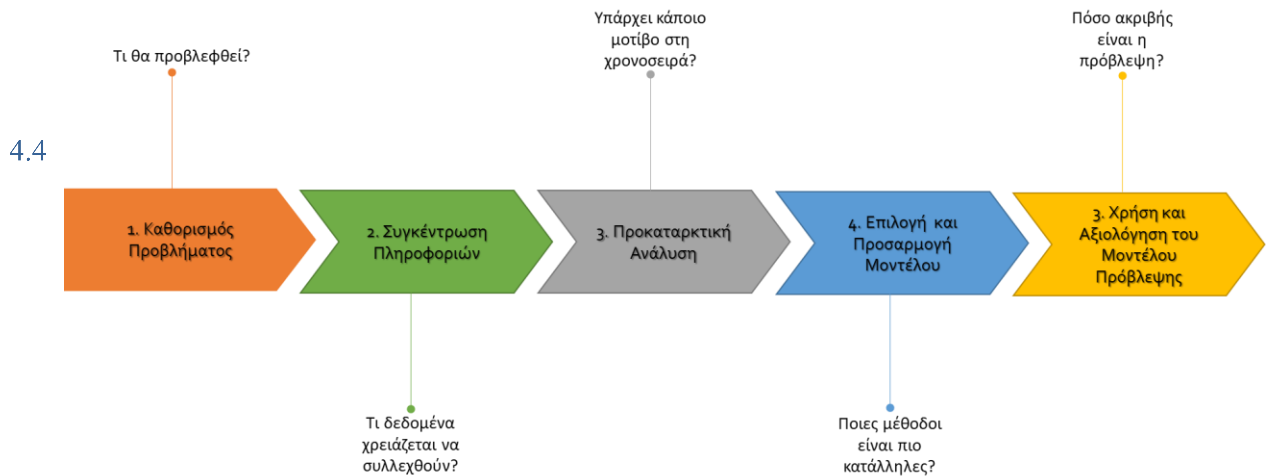
Η εταιρεία για την οποία πραγματοποιήθηκε η μελέτη περίπτωσης (case study) είναι μια ελληνική εταιρεία που δραστηριοποιείται στο χώρο των σνακ, διαθέτοντας πολλά εργοστάσια στην Ελλάδα και πάνω από 1000 εργαζόμενους. Η εταιρεία διαθέτει τμήματα πωλήσεων σε μεγάλες πόλεις, ενώ στην υπόλοιπη Ελλάδα διακινεί τα προϊόντα της μέσω ενός δικτύου τοπικών συνεργατών.

Οι στήλες που περιλαμβάνει το αρχείο που αντλήθηκε από τα συστήματα της εταιρείας είναι οι εξής:

- Έτος
- Μήνας
- Sales: Ποσότητα προϊόντων
- Mother Code
- Code: Κωδικοί προϊόντων
- Ομάδα Brand
- Brand: Εμπορικές επωνυμίες προϊόντων
- Πελάτης: Οι πελάτες στους οποίους πωλήθηκαν τα προϊόντα.

Τα δεδομένα πωλήσεων από το 2017-2019 αφορούν μια συγκεκριμένη κατηγορία προϊόντων και λόγω της εμπιστευτικότητας των δεδομένων έχουν γίνει οι απαιτούμενες παραλλαγές στις ονομασίες. Το dataset περιλαμβάνει 220 κωδικούς προϊόντων, 171 κωδικούς τύπου «mother», 40 διαφορετικά brand, 28 ομάδες brand. Οι πωλήσεις αυτών των προϊόντων έχουν γίνει σε 76 μοναδικούς πελάτες.

Ανάλυση Μεθοδολογίας



Εικόνα 4.1 Βασικά στάδια της πειραματικής διαδικασίας

Όπως είδαμε και στην ενότητα της θεωρίας, η διαδικασία παραγωγής και αξιολόγησης προβλέψεων περνάει από πέντε βασικά στάδια. Σε αυτή την παράγραφο, θα αναλυθεί η πειραματική διαδικασία που διεξάχθηκε βασισμένη σε αυτά τα στάδια.

Στην πρώτη φάση της παρούσας διπλωματικής, οι χρονοσειρές που ορίζουμε προς πρόβλεψη είναι η **ζήτηση ανά brand**. Ο στόχος είναι η παραγωγή μηνιαίων προβλέψεων για το επόμενο εξάμηνο, δηλαδή 6 προβλέψεις. Για να επιτευχθεί αυτό θα αξιοποιήσουμε δυο βασικές τεχνικές πρόβλεψης. Αρχικά θα μελετήσουμε την απόδοση κάποιων βασικών στατιστικών μοντέλων και έπειτα θα δοκιμάσουμε κάποιες μεθόδους μηχανικής μάθησης όπως κάποιο νευρωνικό δίκτυο με ένα δέντρο απόφασης με σκοπό την επιλογή της μεθόδου που θα επιφέρει την πιο ακριβή πρόβλεψη.

4.4.1

Απαιτούμενες βιβλιοθήκες

Στα πλαίσια της συγκεκριμένης διπλωματικής εργασίας χρησιμοποιηθούν ενδεικτικά οι ακόλουθες βιβλιοθήκες της R

- `forecast`: Η συγκεκριμένη βιβλιοθήκη περιέχει αρκετά στατιστικά μοντέλα πρόβλεψης χρονοσειρών τα οποία μπορούν να χρησιμοποιηθούν για γρήγορη και εύκολη παραγωγή προβλέψεων
- `ggplot2`: Οπτικοποίηση δεδομένων
- `plyr`: Συνάθροιση δεδομένων ως προς κάποια κλειδιά
- `RSNNS`: Εφαρμογή του αλγορίθμου MLP

Διαχείριση κενών και αρνητικών τιμών

Τα δεδομένα αποτελούν την «πρώτη ύλη» για τη διαμόρφωση των προβλέψεων. Η ανάλυση των δεδομένων μπορεί να είναι σε ορισμένες περιπτώσεις μια δύσκολη και χρονοβόρα διαδικασία και για αυτό θα πρέπει να της δίνεται ιδιαίτερη βαρύτητα, ώστε
4.4.2 να αντιμετωπίζονται πιθανά προβλήματα που σχετίζονται με τη διαθεσιμότητα αλλά και την ακρίβεια των δεδομένων. Η προ επεξεργασία έχει πραγματοποιηθεί στο excel και έχουμε αθροίσει τις πωλήσεις ανά brand. Υπάρχουν περιπτώσεις όπου λείπουν κάποιες τιμές οι οποίες δημιουργούν προβλήματα στην εφαρμογή των περισσότερων στατιστικών μεθόδων πρόβλεψης και επομένως χρήζουν άμεσης αντιμετώπισης. Εν προκειμένω, δυο χρονοσειρές εμφανίζουν **κενές τιμές** για τα έτη 2017 και 2018 διότι πρόκειται για νέα προϊόντα που λάνσαρε η εταιρεία. Λόγω της έλλειψης, λοιπόν, ιστορικών δεδομένων, επιλέχθηκε να αφαιρεθούν από το dataset και να μη γίνει κάποια διαδικασία για τη συμπλήρωσή τους. Επίσης, παρατηρούμε ότι μια χρονοσειρά παρουσιάζει αρκετές αρνητικές τιμές με την πάροδο των μηνών, γεγονός που υποδεικνύει ότι αυτό το brand οδεύει προς απόσυρση οπότε και αυτό θα αποκλειστεί από την περεταίρω ανάλυσή μας.

Οι **αρνητικές τιμές** που υπάρχουν σε αρκετές χρονοσειρές για τις πωλήσεις δείχνουν τις αντίστοιχες επιστροφές προϊόντων και για τις ανάγκες της συγκεκριμένης εργασίας θα θεωρηθούν μηδενικές. Μέσω της κατάλληλης εντολής, οποιοδήποτε αρνητική τιμή πήρε την τιμή 0.

Η εισαγωγή των δεδομένων θα γίνει μέσω της εντολής read του αρχείου csv που περιέχει μια στήλη για τη χρονιά, μια για τον κάθε μήνα και ακολουθούν οι στήλες που αφορούν το κάθε brand μαζί με τις επικεφαλίδες τους.

Show 50 entries

Search:

	Year	Month	Alexandros	Alkinoos	Alkiviadis.1	Alkiviadis.2	Anaxagoras.1	Anaxagoras.2	Anaxagoras.3	Arhiloxos
1	2017	1		119360	8327	2057	13053	53670	6086	13837
2	2017	2		163697	9272	2399	15730	68123	10361	14108
3	2017	3		214651	11719	3577	21400	77012	10907	24757
4	2017	4		174806	9304	1890	18188	70964	8815	21336
5	2017	5		161635	9066	1668	27772	61544	10528	13328
6	2017	6		186851	9609	386	24235	98493	11506	15478
7	2017	7		187777	10135	38	21083	95163	8373	21693
8	2017	8		198802	11891	0	25819	127010	11089	61573
9	2017	9		156209	10594	0	18432	71996	12966	16566
10	2017	10		175551	12245	3223	15070	53266	12988	16262
11	2017	11		132591	11222	3776	15299	58336	9248	13814
12	2017	12		122599	8596	2461	11009	45451	8782	13902
13	2018	1		108078	9110	2725	9257	51167	8813	12119
14	2018	2		146380	9326	2868	12891	70649	9555	11915
15	2018	3		188660	11482	2899	15617	77333	10202	17150
16	2018	4		149311	9511	1767	14034	71054	8128	13248
17	2018	5	8731	154482	9262	1733	10605	75517	14118	13579
18	2018	6	22098	162074	9609	872	1135	79766	11047	17601
19	2018	7	27884	227677	11137	168	281	103713	9864	18635
20	2018	8	24846	225340	13014	20	-184	118409	9044	62169
21	2018	9	14608	161089	10078	131	-36	68911	10125	19579
22	2018	10	10888	187750	12455	3733	-44	66382	13406	11663
23	2018	11	8959	161713	11870	2802	-39	52421	9479	10262

Εικόνα 4.2 Εισαγωγή δεδομένων στην R

Κατά την ανάπτυξη ενός μοντέλου προβλέψεων επιλέγεται ένα τμήμα των διαθέσιμων δεδομένων για εκμάθηση του μοντέλου και ένα τμήμα για την αξιολόγηση του μοντέλου σε πραγματικές συνθήκες. Συνήθως το ποσοστό διαχωρισμού είναι 80% των δεδομένων για το σύνολο εκπαίδευσης και 20% για το σύνολο δοκιμών. Η μεθοδολογία που ακολουθήθηκε στην παρούσα διπλωματική είναι ο διαχωρισμός του δείγματος σε τρία σύνολα. Αρχικά απομονώσαμε τελείως τις τελευταίες 4 παρατηρήσεις, το σύνολο αξιολόγησης, που θα προσομοίωναν τα πραγματικά μελλοντικά δεδομένα. Στο πρώτο βήμα, χρησιμοποιήσαμε το σύνολο εκπαίδευσης που περιλαμβάνει τους 28 πρώτους μήνες των διαθέσιμων δεδομένων και το σύνολο δοκιμών διάρκειας 4 μηνών προκειμένου να αξιολογήσουμε ποιο μοντέλο παράγει τις πιο ακριβείς προβλέψεις ανά χρονοσειρά βάσει του συμμετρικού μέσου απόλυτου ποσοστιαίου σφάλματος - sMAPE που αναλύθηκε σε προηγούμενη ενότητα.

Έπειτα, ενοποιήσαμε αυτά τα δεδομένα διάρκειας 32 μηνών και συγκρίναμε το επιλεγμένο μοντέλο πρόβλεψης με το σύνολο αξιολόγησης διάρκειας 4 μηνών προκειμένου να διαπιστώσουμε ένα το επιλεγμένο μοντέλο εξακολουθούσε να έχει την ίδια απόδοση.



Εικόνα 4.3 Διαχωρισμός του συνόλου δεδομένων

Οπτικοποίηση δεδομένων

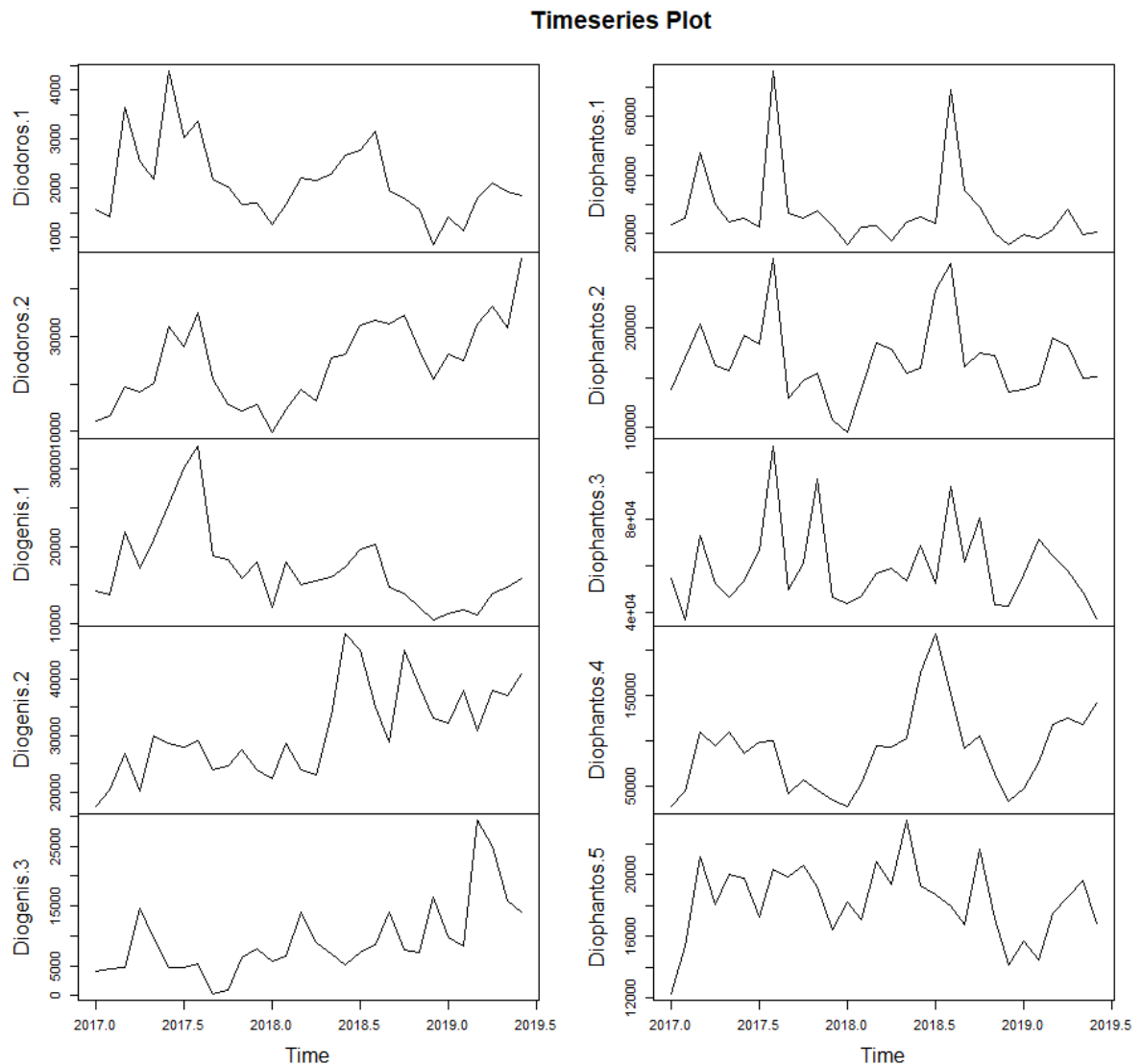
Εφόσον αντιμετωπίστηκαν οι κενές και αρνητικές τιμές, πλέον μπορούν να απεικονιστούν τα δεδομένα σε γραφήματα προκειμένου να αναγνωριστούν μοτίβα που δεν είναι εμφανή όταν τα δεδομένα είναι σε μορφή πίνακα. Η οπτικοποίηση των 4.4.3 δεδομένων είναι πολύ χρήσιμη στην επιστήμη των προβλέψεων καθώς βοηθάει τον αναλυτή να προσεγγίσει καλύτερα το πρόβλημα.

Μια χρονοσειρά μπορεί να εκφραστεί στην R ως δομημένο αντικείμενο με την έννοια του χρόνου, χρησιμοποιώντας τη συνάρτηση `ts`, η οποία μετατρέπει την ακολουθία αριθμών σε χρονοσειρά. Οι χρονοσειρές που χρησιμοποιούμε δηλώνουμε ότι έχουν συχνότητα 12 μήνες και ξεκινάνε το 2017.

Η γλώσσα R παρέχει μια πληθώρα πακέτων για προηγμένες απεικονίσεις. Αυτά τα μοτίβα, αφορούν τις τέσσερις βασικές συνιστώσες των χρονοσειρών: τάση, εποχιακότητα, κυκλικότητα και τυχαιότητα. Η τάση μπορεί να είναι ανοδική, πτωτική ή μηδενική. Η κυκλικότητα, όπως αναφέρθηκε και σε προηγούμενη ενότητα, αφορά κυρίως οικονομικές συνθήκες και δεν θα αναλυθεί περεταίρω στην παρούσα διπλωματική. Η εποχιακότητα, που αντιπροσωπεύει τις περιοδικές διακυμάνσεις που έχουν σταθερό μήκος, θα αποτελέσει αντικείμενο προς μελέτη παρακάτω.

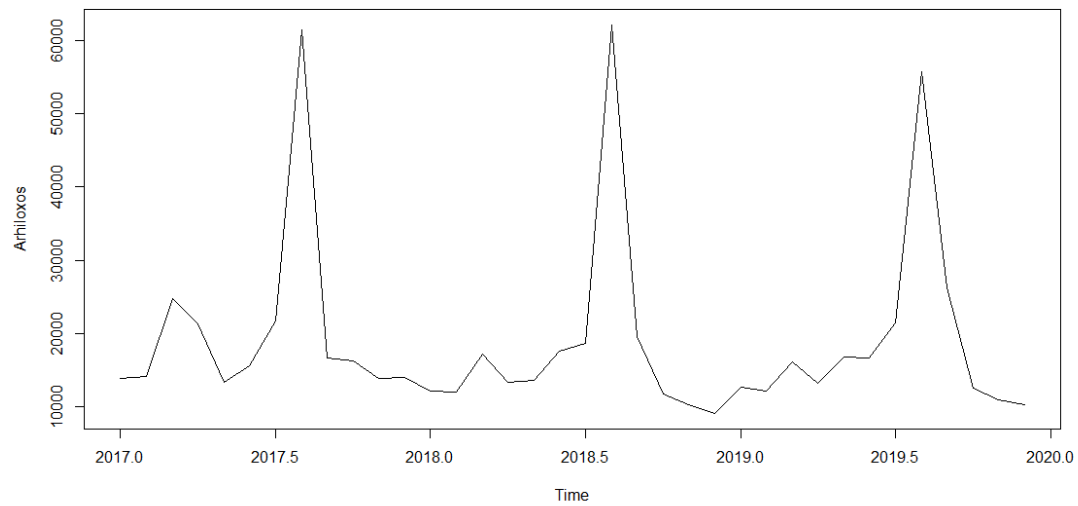
Κάθε χρονοσειρά παρουσιάζει τα δικά της χαρακτηριστικά. Όπως παρατηρείται, κάποιες εμφανίζουν έντονη εποχιακότητα καθώς μπορεί να αφορούν προϊόντα που έχουν ζήτηση συγκεκριμένες εποχές του χρόνου, όπως τα παγωτά, ενώ κάποιες άλλες εμφανίζουν τάση αλλά και έντονες διακυμάνσεις, δηλαδή τυχαιότητα.

Στην R, επιλέξαμε να δούμε αρκετές χρονοσειρές μαζί για να παρατηρήσουμε ομοιότητες και διαφορές αλλά μελετήσαμε και κάθε χρονοσειρά μόνη της.

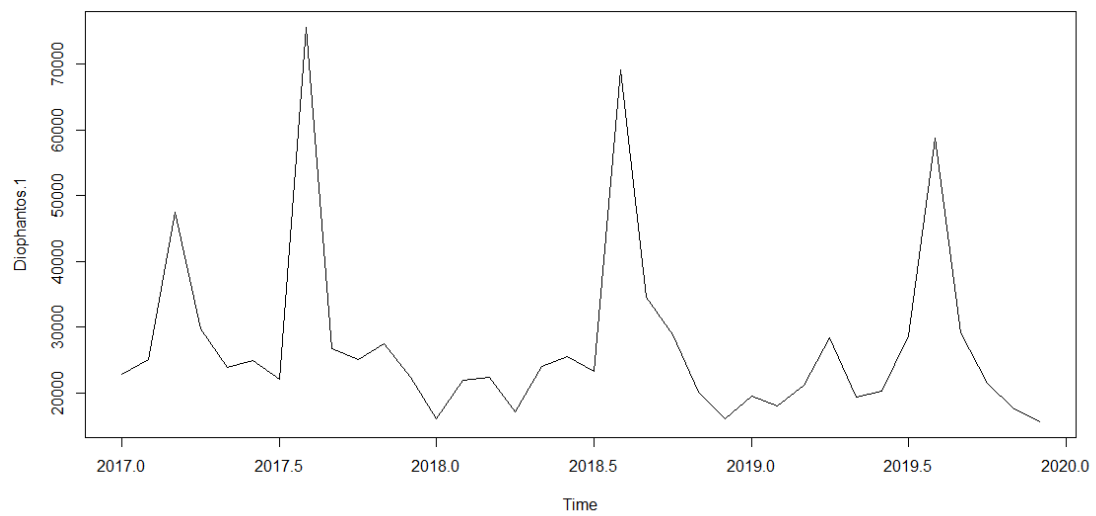


Εικόνα 4.4 Απεικόνιση πολλαπλών χρονοσειρών

Στις παρακάτω χρονοσειρές είναι εμφανής η εποχιακότητα και συγκεκριμένα η αύξηση των πωλήσεων κατά τους καλοκαιρινούς μήνες. Στη δεύτερη χρονοσειρά, εκτός από την εποχιακότητα παρατηρείται και μια ελαφριά φθίνουσα τάση.

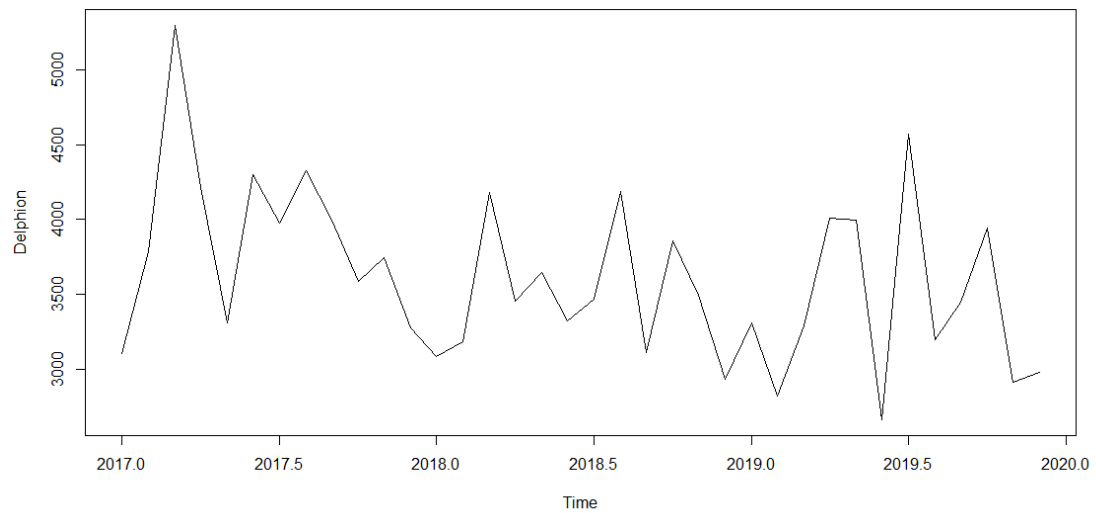


Εικόνα 4.5 Χρονοσειρά με εποχιακότητα

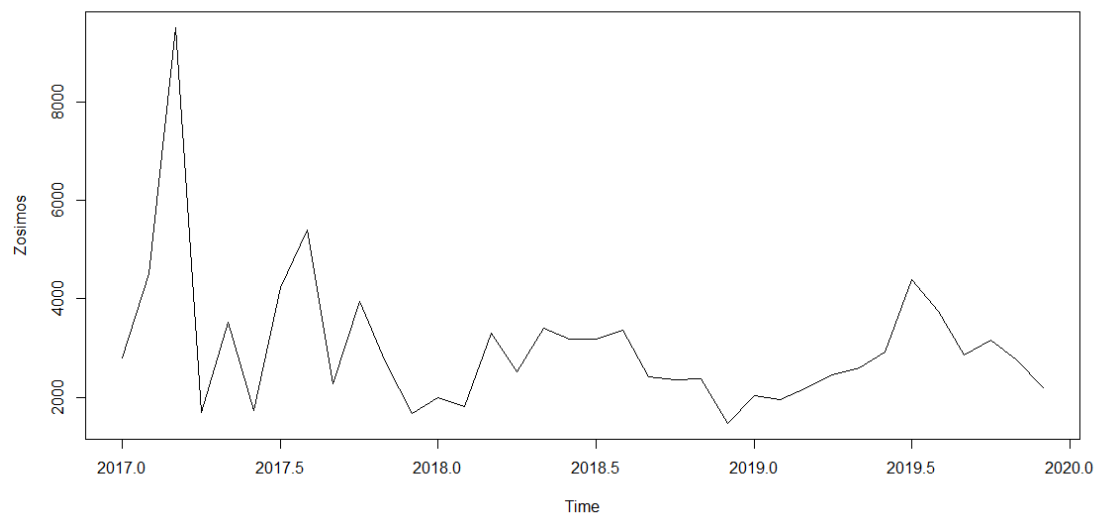


Εικόνα 4.6 Χρονοσειρά με εποχιακότητα και τάση

Στις παρακάτω χρονοσειρές παρατηρείται μια πτωτική τάση ενώ η δεύτερη χρονοσειρά εμφανίζει και τυχαιότητα αφού



Εικόνα 4.7 Χρονοσειρά με πτωτική τάση



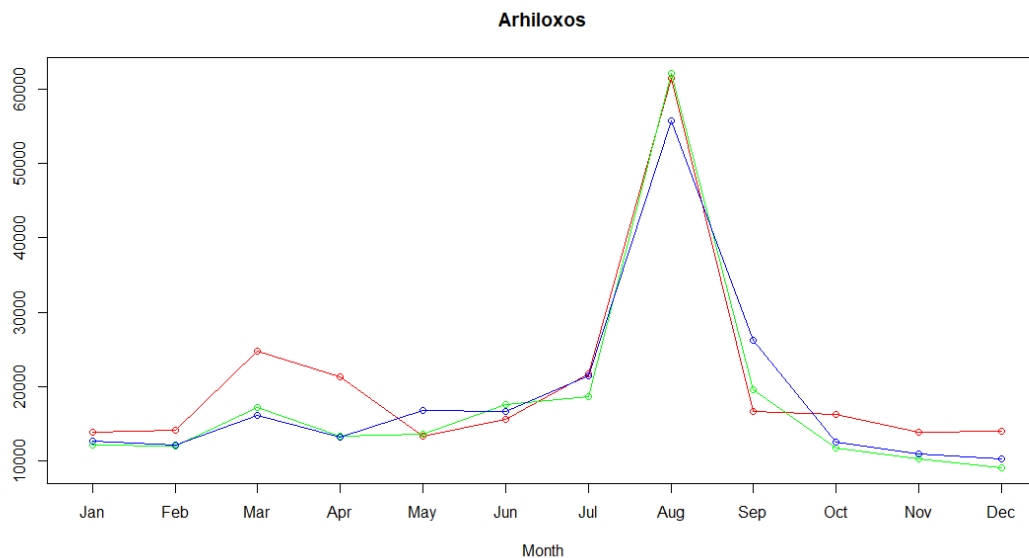
Εικόνα 4.8 Χρονοσειρά με πτωτική τάση και τυχαιότητα

4.4.3.1 Διαγράμματα εποχιακότητας

Αν και τα παραπάνω γραφήματα μας βοηθούν να αρχίσουμε να αντιλαμβανόμαστε πώς συμπεριφέρεται η εκάστοτε χρονοσειρά, θα χρησιμοποιήσουμε και την εντολή `ggseasonplot` για το διάγραμμα εποχιακότητας για να κατανοήσουμε πλήρως την εποχιακή συμπεριφορά των χρονοσειρών μας. Η εντολή `ggseasonplot` χωρίζει τη χρονοσειρά σε ισομήκη κομμάτια μήκους όσο η συχνότητα που θεωρούμε για τα δεδομένα μας, οπότε στην ανάλυση μας θα είναι χωρισμένη ανά έτος, 12 μήνες. Η γνώση της εποχιακότητας των πωλήσεων είναι ύψιστης σημασίας για τις αποφάσεις που λαμβάνουν τα διοικητικά στελέχη σχετικά με τον προγραμματισμό της παραγωγής. Γνωρίζοντας αυτή την πληροφορία, είναι δυνατή η πρόβλεψη μεταβολών στις χρονοσειρές που θα οδηγήσουν και σε αντίστοιχες αποφάσεις. Επίσης, η αναγνώριση της εποχιακότητας είναι απαραίτητη για τον προσδιορισμό του ιδανικού μοντέλου πρόβλεψης.

Για τον προσδιορισμό της εποχιακότητας, ο αναλυτής έχει τις παρακάτω δύο επιλογές:

- Χρήση μεθόδου που λαμβάνει υπόψη την εποχιακότητα
- Αποσύνθεση της χρονοσειράς, δημιουργία πρόβλεψης και έπειτα εποχικοποίηση της παραχθείσας πρόβλεψης



Εικόνα 4.9 Απεικόνιση χρονοσειράς σε διάγραμμα εποχιακότητας

Επιλογή μεθόδων πρόβλεψης

Σε προηγούμενη ενότητα, αναφέρθηκαν διάφορες χρήσιμες και δημοφιλείς τεχνικές πρόβλεψης χρονοσειρών. Το επόμενο σημαντικό ζήτημα είναι η εφαρμογή κάποιας μεθόδου ώστε να είναι δυνατή η πρόβλεψη της χρονοσειράς.

4.4.4 Στη συγκεκριμένη διπλωματική καλούμαστε να αντιμετωπίσουμε το πρόβλημα της προετοιμασίας προβλέψεων για τους επόμενους 4 μήνες για 37 διαφορετικά προϊόντα της γραμμής παραγωγής. Παρατηρώντας τα εποχιακά διαγράμματα εξήχθη το συμπέρασμα ότι κάποιες χρονοσειρές παρουσιάζουν εποχιακότητα, έντονη και μη και κάποιες άλλες όχι. Συνεπώς αποφασίσαμε να χρησιμοποιήσουμε κάποιες αυτόματες συναρτήσεις πρόβλεψης που δεν έχουν ως προ-απαιτούμενο την αποσύνθεση της χρονοσειράς ώστε να είναι αποδοτικές. Συγκεκριμένα εφαρμόσαμε τις μεθόδους SNaive, Εξομάλυνση ETS, ARIMA και Theta. Όπως αναλύσαμε σε προηγούμενα ενότητα η μέθοδος Seasonal Naïve(SNaive) επιστρέφει σαν πρόβλεψη την τιμή της τελευταίας παρατήρησης της αντίστοιχης εποχιακής περιόδου. Οι μέθοδοι ETS και ARIMA ανάλογα τα δεδομένα αποφασίζουν αυτόματα ποιο συγκεκριμένο μοντέλο θα εφαρμοστεί και η συνάρτηση `thetaf()`, που χρησιμοποιήθηκε για την εφαρμογή της μεθόδου Theta, ελέγχει αυτόματα τις χρονοσειρές για ύπαρξη εποχιακότητας και σε περίπτωση που εμφανίζεται το εποχικό στοιχείο εξαλείφεται με πολλαπλασιαστική αποσύνθεση προτού εφαρμοστεί η μέθοδος.

Πρόκειται για μεθόδους αρκετά διαφορετικές μεταξύ τους και παρουσιάζει αρκετά μεγάλο ενδιαφέρον η απόδοσή τους σε μια γκάμα χρονοσειρών με διαφορετικά χαρακτηριστικά.

Αξιολόγηση Μεθόδων

Για να αξιολογήσουμε την ακρίβεια των προβλέψεων μεταξύ διαφορετικών μοντέλων που εφαρμόσαμε στις χρονοσειρές μας, χρησιμοποιήσαμε το στατιστικό δείκτη το **4.5 συμμετρικό μέσο απόλυτο ποσοστιαίο σφάλμα (Symmetric Mean Absolute Percentage Error) – sMAPE**. Επιλέξαμε ένα σφάλμα που να είναι σε ποσοστιαία μορφή προκειμένου να μπορέσουμε να συγκρίνουμε την ακρίβεια μιας μεθόδου στις διαθέσιμες χρονοσειρές, όπου η καθεμιά έχει διαφορετικό επίπεδο μέσης τιμής. Επίσης, επιλέχθηκε ο συγκεκριμένος δείκτης, που αποτελεί μια παραλλαγή το μέσου απόλυτου ποσοστιαίου σφάλματος, επειδή μπορεί να εφαρμοσθεί και σε χρονοσειρές όπου η ζήτηση παίρνει και την τιμή μηδέν και όπως αναλύθηκε παραπάνω τις αρνητικές τιμές τις μηδενίσαμε.

Αποτελέσματα στατιστικών μεθόδων

4.5.1 Μέσω ενός βρόχου επανάληψης, η κάθε χρονοσειρά παρήγαγε προβλέψεις βάσει του συνόλου εκπαίδευσης (28 μήνες) και με τις 4 μεθόδους SNaive, ETS, ARIMA, Theta. Έπειτα, ο αλγόριθμος υπολόγισε το σφάλμα πρόβλεψης όπου σύγκρινε τα ιστορικά δεδομένα που αποθηκεύσαμε στο σύνολο δοκιμών(4 μήνες) με τις προβλέψεις που έχουν προκύψει από τις μεθόδους που επιλέξαμε. Στην R μέσω του \$mean εξήγαμε τις προβλέψεις για τον ορίζοντα που είχαμε ορίσει.

	SNaive	ETS	Arima	Theta
Alkinoos	11.036	16.364	9.249	13.854
Alkiviadis.1	7.521	10.301	4.078	10.012
Alkiviadis.2	20.77	20.922	20.77	125.844
Anaxagoras.2	4.035	9.918	4.035	29.796
Anaxagoras.3	17.045	9.542	17.045	9.541
Arhiloxos	12.761	20.479	18.201	11.314
Arion	21.579	47.97	21.579	49.019
Aristeas	11.232	13.973	14.479	10.022
Aristomaxos	17.991	53.998	44.357	60.726
Aristotelis.1	10.345	26.056	10.345	23.786
Aristotelis.2	9.198	5.644	5.641	6.927
Arkas	14.498	20.933	6.937	21.65
Asklipios	7.128	6.104	7.128	5.878
Avdiros	13.586	26.575	6.118	16.733
Dedalos	25.155	29.031	25.373	28.591

Delphion	21.338	18.948	19.031	18.734
Dimokritos.1	20.012	11.083	32.597	46.143
Dimokritos.3	43.984	53.904	51.907	58.039
Dimosthenis	9.964	21.24	9.964	21.206
Diodoros.1	17.949	18.565	13.414	20.875
Diodoros.2	45.95	31.799	16.299	30.419
Diogenis.1	7.388	28.942	26.836	31.235
Diogenis.2	18.664	11.472	13.471	18.628
Diogenis.3	79.744	11.732	12.086	33.918
Diophantos.1	20.37	35.465	20.37	28.51
Diophantos.2	3.351	24.257	8.485	24.236
Diophantos.3	31.136	38.444	31.136	38.366
Diophantos.4	14.77	17.09	17.904	17.113
Diophantos.5	11.494	8.489	7.272	8.447
Dymas	27.412	16.724	48.76	16.578
Epafos	12.813	21.572	12.077	22.298
Periklis	23.583	22.409	22.409	22.342
Sofoklis.1	22.328	8.744	14.112	44.125
Sofoklis.2	8.616	33.985	8.616	34.569
Sokratis	10.617	16.18	10.617	14.089
Zosimos	19.521	39.475	33.65	45.815
AVERAGE	18.747	22.454	17.954	28.316

Πίνακας 4.1 Σφάλματα στατιστικών μεθόδων

Παρατηρούμε, ότι η μέθοδος με το μικρότερο μέσο σφάλμα είναι η ARIMA με ποσοστό 17,954% ενώ με μικρή διαφορά ακολουθεί η απλοϊκή μέθοδος SNaive με 18,747%.

Το επόμενο βήμα του αλγορίθμου ήταν, για την κάθε χρονοσειρά, η μέθοδος που εμφάνισε το μικρότερο σφάλμα sMAPE να χρησιμοποιηθεί και για την τελική πρόβλεψη της χρονοσειράς με βραχυπρόθεσμο ορίζοντα 4 μηνών.

Η απόδοση κάθε μεθόδου καθώς και της τελικής πρόβλεψης εμφανίζονται στον παρακάτω πίνακα.

	Snaive	ETS	Arima	Theta	Final Statistical Selection
Alkinoos	9.652	22.648	9.652	8.742	9.652
Alkiviadis.1	4.739	13.638	5.092	5.396	5.092
Alkiviadis.2	71.058	70.824	71.058	131.91	71.058
Anaxagoras.2	26.914	21.324	26.914	23.645	26.914
Anaxagoras.3	7.184	15.011	7.184	14.983	14.983

	Snaive	ETS	Arima	Theta	Final Statistical Selection
Arhiloxos	13.533	16.084	28.081	15.849	15.849
Arion	4.233	19.837	4.233	19.346	4.233
Aristeas	18.059	9.879	18.059	22.437	22.437
Aristomaxos	16.97	12.158	13.429	42.607	16.97
Aristotelis.1	8.747	43.517	7.221	44.398	8.747
Aristotelis.2	13.875	10.55	10.55	10.757	10.55
Arkass	3.741	55.54	23.832	54.321	23.832
Asklipios	11.624	11.465	11.624	10.474	10.474
Avdiros	13.576	23.577	5.695	5.281	5.695
Dedalos	10.139	18.827	12.371	19.177	10.139
Delphion	8.022	14.002	13.999	10.893	10.893
Dimokritos.1	19.77	11.023	12.617	55.104	11.023
Dimokritos.3	36.687	45.078	44.399	40.874	36.687
Dimosthenis	5.34	6.888	2.753	2.756	5.34
Diodoros.1	20.098	68.303	21.552	67.16	21.552
Diodoros.2	8.904	66.58	49.982	63.914	49.982
Diogenis.1	5.883	41.449	37.684	40.033	5.883
Diogenis.2	15.418	12.903	8.707	18.622	12.903
Diogenis.3	33.643	6.787	6.38	17.337	6.787
Diophantos.1	15.257	14.005	12.386	17.588	15.257
Diophantos.2	12.339	14.843	12.339	16.123	12.339
Diophantos.3	13.818	13.966	13.818	14.46	13.818
Diophantos.4	18.94	4.839	11.384	71.437	18.94
Diophantos.5	5.75	15.484	8.929	14.33	8.929
Dymas	52.099	10.657	12.37	10.645	10.645
Epafos	6.287	18.402	14.448	9.536	14.448
Periklis	12.472	5.058	5.06	5.082	5.082
Sofoklis.1	11.887	8.801	12.704	9.509	8.801
Sofoklis.2	8.754	7.331	8.754	52.154	8.754
Sokratis	10.663	2.372	10.663	7.523	10.663
Zosimos	25.138	24.069	13.333	18.93	25.138
AVERAGE	16.145	21.603	16.368	27.593	15.847

Πίνακας 4.2 Σύγκριση σφαλμάτων στατιστικών μεθόδων με μέθοδο αυτόματης επιλογής

Σε αυτό το σημείο θα πρέπει να αναφέρουμε κάποιες γενικές παρατηρήσεις σχετικά με τα αποτελέσματα που παρατέθηκαν στους παραπάνω πίνακες.

- Σε αρκετές χρονοσειρές η μέθοδος seasonal Naïve, αν και απλοϊκή μέθοδος, φάνηκε να παρουσιάζει μεγαλύτερη ακρίβεια σε σχέση με τις υπόλοιπες μεθόδους, γεγονός που οφείλεται στην έντονη εποχιακότητα κάποιων χρονοσειρών(π.χ. Αριστόμαχος) αφού δεν τις αποσυνθέσαμε προτού εφαρμόσουμε τις μεθόδους πρόβλεψης.
- Επίσης, με βάση τα τελικά δεδομένα, παρατηρούμε ότι η μέθοδος seasonal Naïve έχει το μικρότερο μέσο σφάλμα 16,145% συγκριτικά με τις υπόλοιπες 3 μεθόδους.
- Παρατηρήθηκε σε κάποιες χρονοσειρές ότι αν και το σφάλμα πρόβλεψης στα αρχικά δεδομένα ήταν αρκετά μικρό(<10%), κατά την τελική πρόβλεψη το σφάλμα αυξήθηκε γεγονός που μπορεί να οφείλεται στην τυχαιότητα που μπορεί να υπήρχε στους μήνες που κληθήκαμε να προβλέψουμε. Για παράδειγμα, η χρονοσειρά «Αναξαγόρας.3» στα αρχικά δεδομένα είχε σφάλμα 9,541% ενώ κατά την τελική πρόβλεψη το σφάλμα ήταν 14,983%.
- Η επιλογή της βέλτιστης μεθόδου ανά χρονοσειρά είχε ως αποτέλεσμα το μικρότερο μέσο σφάλμα, δηλαδή 15,847%. Αν είχαμε επιλέξει τη μέθοδο που είχε το μικρότερο μέσο σφάλμα στα αρχικά δεδομένα για να την εφαρμόσουμε στην τελική πρόβλεψη, θα είχαμε ως αποτέλεσμα το σφάλμα της μεθόδου ARIMA που είναι 16,368%. Συνεπώς, με την αυτοματοποίηση του αλγορίθμου επιτεύχθηκε μείωση του σφάλματος κατά 0,521%.

Σε επόμενο στάδιο, εξετάσαμε την περίπτωση συνδυασμού κάποιων μεθόδων και συγκεκριμένα των μεθόδων sNaïve, ETS, και ARIMA λόγω του ότι παρουσίασαν το μικρότερο μέσο σφάλμα κατά την πρόβλεψη των αρχικών δεδομένων. Σε κάποιες περιπτώσεις, ακόμα και να αναγνωριστούν σωστά οι διάφοροι παράγοντες, είναι δύσκολη η εύρεση μίας μεθόδου αποκλειστικά που να ικανοποιεί πλήρως όλες τις απαιτήσεις του προβλήματος για την πρόβλεψη. Παρατηρήθηκε ότι ένας τρόπος αύξησης της ακρίβειας των προβλέψεων, είναι ο συνδυασμός διαφορετικών μεθόδων πρόβλεψης. Διαφορετικές μέθοδοι πρόβλεψης, εφαρμοζόμενες στις ίδιες χρονοσειρές, παράγουν διαφορετικά αποτελέσματα καθώς η κάθε πρόβλεψη παρέχει και διαφορετική πληροφορία. Ο συνδυασμός τους παράγει προβλέψεις, αξιοποιώντας

περισσότερη πληροφορία στην ίδια τιμή, γεγονός που ενισχύει την ακρίβεια της πρόβλεψης καθώς προσεγγίζεται καλύτερα η πραγματικότητα.

Η μέθοδος που ακολουθήθηκε για την υλοποίηση του εν λόγω συνδυασμού είναι ο απλός μέσος όρος όλων των προβλέψεων των μεθόδων. Μελέτες που έχουν πραγματοποιηθεί έχουν αποδείξει ότι ο υπολογισμός του απλού μέσου όρου προβλέψεων οδηγεί σε αποτελέσματα το ίδιο ικανοποιητικά με αυτά πολύπλοκων τεχνικών συνδυασμού.

Παρατηρούμε πως στις υπό μελέτη χρονοσειρές ο συνδυασμός μεθόδων επέφερε μείωση του σφάλματος sMAPE κατά 2,586% αν και σε κάποιες χρονοσειρές ο συνδυασμός των μεθόδων έχει πιο μεγάλα σφάλματα από την τελική μεμονωμένη μέθοδο, όπως για παράδειγμα στη χρονοσειρά «Αριστοτέλης.1», όπου το σφάλμα της Combo μεθόδου είναι 14,746% ενώ το σφάλμα της τελικής μεμονωμένης μεθόδου είναι 8,747%. Στις περισσότερες χρονοσειρές, ωστόσο, ο συνδυασμός των μεθόδων παρουσιάζει μεγαλύτερη ακρίβεια με το μικρότερο σφάλμα να είναι μόλις 3,1% στη χρονοσειρά «Σοφοκλής.1»

	Combo (sNaive,ETS,Arima)	Final Statistical Selection
Alkinoos	8.067	9.652
Alkiviadis.1	7.113	5.092
Alkiviadis.2	70.98	71.058
Anaxagoras.2	24.937	26.914
Anaxagoras.3	7.17	14.983
Arhiloxos	14.111	15.849
Arion	6.691	4.233
Aristeas	15.36	22.437
Aristomaxos	9.902	16.97
Aristotelis.1	14.746	8.747
Aristotelis.2	8.077	10.55
Arkas	18.909	23.832
Asklipios	11.562	10.474
Avdiros	7.13	5.695
Dedalos	9.171	10.139
Delphion	10.245	10.893
Dimokritos.1	14.338	11.023
Dimokritos.3	19.633	36.687
Dimosthenis	4.748	5.34
Diodoros.1	32.456	21.552
Diodoros.2	42.717	49.982
Diogenis.1	6.147	5.883

	Combo (sNaive,ETS,Arima)	Final Statistical Selection
Diogenis.2	4.17	12.903
Diogenis.3	7.35	6.787
Diophantos.1	10.623	15.257
Diophantos.2	13.166	12.339
Diophantos.3	6.384	13.818
Diophantos.4	10.649	18.94
Diophantos.5	5.971	8.929
Dymas	14.261	10.645
Epafos	5.427	14.448
Periklis	7.467	5.082
Sofoklis.1	3.1	8.801
Sofoklis.2	8.216	8.754
Sokratis	6.981	10.663
Zosimos	9.404	25.138
AVERAGE	13.261	15.847

Πίνακας 4.3 Σύγκριση συνδυασμού μεθόδων(Combo) με μέθοδο αυτόματης επιλογής

Αποτελέσματα μεθόδων μηχανικής μάθησης

4.5.2

Έπειτα από την υλοποίηση των στατιστικών μεθόδων, εξετάστηκαν διάφορες προσεγγίσεις βασισμένες στα μοντέλα μηχανικής μάθησης και πιο συγκεκριμένα στα πολυεπίπεδα αντίληπτρα(MLP). Όπως προαναφέρθηκε, τα μοντέλα μηχανικής μάθησης είναι πιο αποδοτικά όταν εφαρμόζονται σε μεγάλο όγκο διαθέσιμων δεδομένων. Προκειμένου να ξεπεράσουμε αυτό τον σκόπελο, εκπαιδεύσαμε ένα ενιαίο νευρωνικό δίκτυο από όλα τα δεδομένα όλων των χρονοσειρών, και έπειτα βάσει των μοτίβων που ανακάλυψε κλήθηκε να προβλέψει τη ζήτηση για κάθε χρονοσειρά ξεχωριστά. Όσον αφορά στην εκπαίδευση των μοντέλων πραγματοποιήθηκαν διάφορες δοκιμές ώστε να προκύψει ποια είναι η καλύτερη προσέγγιση για τα δεδομένα μας. Οι δοκιμές αφορούσαν κυρίως τη συνάρτηση εκμάθησης και τη συνάρτηση ενεργοποίησης. Ως συνάρτηση εκμάθησης χρησιμοποιήθηκε η ευρέως χρησιμοποιούμενη *Std_Backpropagation* που αναλύθηκε και παραπάνω και η *Rprop* που στη βιβλιογραφία δίνει ικανοποιητικά αποτελέσματα. Σχετικά με τη συνάρτηση ενεργοποίησης δοκιμάστηκε η *Act_Logistic*, η *Act_IdentityPlusBias* και η *Act_Exponential*. Διατηρήσαμε το *size=5* σταθερό για όλα τα μοντέλα. Οι συνδυασμοί που καθόρισαν τα μοντέλα που εξετάσαμε με τα αντίστοιχά τους σφάλματα παρουσιάζονται στον παρακάτω πίνακα.

	Learning Function	Activation Function	sMAPE
Model 1	Std_Backpropagation	Act Logistic	16.354
Model 2	Std_Backpropagation	Act_Exponential	16.340
Model 3	Std_Backpropagation	Act_IdentityPlusBias	17.588
Model 4	Rprop	Act Logistic	17.115
Model 5	Rprop	Act_Exponential	17.966

Πίνακας 4.4 Συγκεντρωτικά σφάλματα μοντέλων MLP

Τα αποτελέσματα του κάθε μοντέλου για την κάθε χρονοσειρά ξεχωριστά παρουσιάζονται στον παρακάτω πίνακα.

	Model 1	Model 2	Model 3	Model 4	Model 5
Alkinoos	6.824	5.199	8.084	7.479	7.058
Alkiviadis.1	10.616	8.357	9.494	9.487	12.435
Alkiviadis.2	64.866	73.577	67.592	72.962	57.911
Anaxagoras.2	10.502	11.895	7.66	10.359	8.362
Anaxagoras.3	14.274	12.805	15.826	10.374	9.878
Arhiloxos	42.236	36.22	42.912	38.553	49.498
Arion	9.324	13.049	15.427	7.402	9.624
Aristeas	5.647	5.175	8.235	7.357	5.453
Aristomaxos	30.436	30.049	35.068	21.415	22.246
Aristotelis.1	7.379	6.863	6.438	13.311	12.339
Aristotelis.2	14.351	15.621	18.096	6.954	9.643
Arkas	11.939	14.548	12.864	17.143	18.632
Asklipios	12.235	10.97	12.273	14.74	16.715
Avdiros	7.957	7.474	8.155	8.244	10.975
Dedalos	7.63	7.861	8.899	10.037	10.503
Delphion	5.643	8.64	7.486	9.588	11.02
Dimokritos.1	22.204	26.224	35.612	18.226	15.82
Dimokritos.3	23.256	13.708	27.326	24.011	32.867
Dimosthenis	6.621	5.313	8.045	12.069	12.654
Diodoros.1	29.362	27.528	30.528	37.781	38.041
Diodoros.2	7.15	7.19	9.948	15.863	14.387
Diogenis.1	13.323	10.208	13.145	12.416	16.091
Diogenis.2	8.251	7.526	9.197	5.63	3.621
Diogenis.3	49.757	51.637	36.821	21.967	26.555
Diophantos.1	37.668	36.302	35.088	42.184	44.364
Diophantos.2	8.549	7.671	10.137	7.006	8.42
Diophantos.3	10.976	12.535	8.59	12.727	15.799
Diophantos.4	24.284	23.418	35.144	40.824	35.954
Diophantos.5	5.564	6.587	5.791	9.533	10.677
Dymas	39.055	38.472	34.716	33.354	34.803
Epafos	11.038	13.235	11.145	14.086	15.516

	Model 1	Model 2	Model 3	Model 4	Model 5
Periklis	5.838	6.114	4.226	4.487	4.794
Sofoklis.1	6.519	6.673	7.442	12.218	12.404
Sofoklis.2	6.841	8.017	12.208	8.481	10.568
Sokratis	4.941	4.845	6.314	3.655	3.734
Zosimos	5.704	6.737	7.227	14.234	17.419
AVERAGE	16.354	16.340	17.588	17.115	17.966

Πίνακας 4.5 Σφάλματα μοντέλων MLP

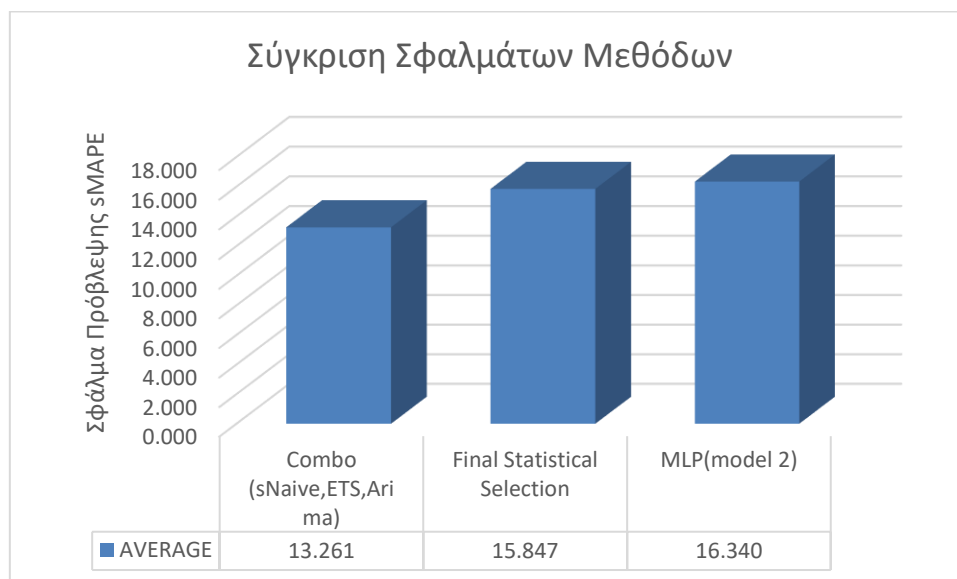
Βάσει αποτελεσμάτων, φαίνεται ότι το μοντέλο που έχει το μικρότερο σφάλμα είναι το μοντέλο 2 με size=5, συνάρτηση εκμάθησης την *Std_Backpropagation* και συνάρτηση ενεργοποίησης την *Act_Exponential*.

Ως τελευταίο βήμα, συγκρίνουμε τις στατιστικές μεθόδους με τα μοντέλα των νευρωνικών δικτύων. Συγκεκριμένα παρατηρούμε ότι η καλύτερη μέθοδος πρόβλεψης συνεχίζει να είναι ο συνδυασμός των μεθόδων sNaive, ETS και ARIMA.

	Combo (sNaive,ETS, Arima)	Final Statistical Selection	MLP (model 2)
Alkinoos	8.067	9.652	5.199
Alkiviadis.1	7.113	5.092	8.357
Alkiviadis.2	70.98	71.058	73.577
Anaxagoras.2	24.937	26.914	11.895
Anaxagoras.3	7.17	14.983	12.805
Arhiloxos	14.111	15.849	36.22
Arion	6.691	4.233	13.049
Aristeas	15.36	22.437	5.175
Aristomaxos	9.902	16.97	30.049
Aristotelis.1	14.746	8.747	6.863
Aristotelis.2	8.077	10.55	15.621
Arkias	18.909	23.832	14.548
Asklipios	11.562	10.474	10.97
Avdiros	7.13	5.695	7.474
Dedalos	9.171	10.139	7.861
Delphion	10.245	10.893	8.64
Dimokritos.1	14.338	11.023	26.224
Dimokritos.3	19.633	36.687	13.708
Dimosthenis	4.748	5.34	5.313
Diodoros.1	32.456	21.552	27.528
Diodoros.2	42.717	49.982	7.19
Diogenis.1	6.147	5.883	10.208

Diogenis.2	4.17	12.903	7.526
Diogenis.3	7.35	6.787	51.637
Diophantos.1	10.623	15.257	36.302
Diophantos.2	13.166	12.339	7.671
Diophantos.3	6.384	13.818	12.535
Diophantos.4	10.649	18.94	23.418
Diophantos.5	5.971	8.929	6.587
Dymas	14.261	10.645	38.472
Epafos	5.427	14.448	13.235
Periklis	7.467	5.082	6.114
Sofoklis.1	3.1	8.801	6.673
Sofoklis.2	8.216	8.754	8.017
Sokratis	6.981	10.663	4.845
Zosimos	9.404	25.138	6.737
AVERAGE	13.261	15.847	16.340

Πίνακας 4.6 Σύγκριση σφαλμάτων όλων των μεθόδων



Εικόνα 4.10 Παρουσίαση Σφαλμάτων στατιστικών μεθόδων, συνδυασμού τους και NN

Συμπεράσματα και Μελλοντικές Προεκτάσεις

Συνοπτικά Συμπεράσματα

5
5.1 Στην παρούσα διπλωματική πραγματοποιήθηκε πρόβλεψη ζήτησης στα brands μιας ελληνικής εταιρίας μεσαίου μεγέθους που δραστηριοποιείται στο χώρο των σνακ. Συγκεκριμένα μελετήθηκαν τα στατιστικά μοντέλα sNaive, ETS , ARIMA και Theta καθώς και ο συνδυασμός τους. Επίσης, πραγματοποιήθηκαν προβλέψεις νευρωνικού δικτύου μέσω του αλγορίθμου MLP. Για την εκπαίδευση του νευρωνικού δικτύου χρησιμοποιήθηκαν όλες οι τιμές των χρονοσειρών σε ένα ενιαίο δίκτυο. Στόχος αυτής της ανάλυσης ήταν ο προσδιορισμός του οφέλους από τη χρήση προηγμένων μοντέλων πρόβλεψης με απώτερο σκοπό τη σύνδεσής τους με το υπάρχον εργαλείο ERP της εταιρείας για να επιτευχθεί αύξηση της ακρίβειας της πρόβλεψης ζήτησης και κατ' επέκταση πιο στοχευμένος προγραμματισμός παραγωγής για την κάλυψη των αναγκών της εταιρείας.

Τα βασικά συμπεράσματα που προκύπτουν από την παραπάνω διαδικασία είναι :

- ο Η κάθε χρονοσειρά παρουσιάζει διαφορετικά χαρακτηριστικά (εποχιακότητα, τάση, κ.λπ) οπότε η αυτόματη επιλογή της καλύτερης μεθόδου βάσει του αρχικού σφάλματος sMAPE μπορεί να επιφέρει αύξηση της ακρίβειας της τελικής πρόβλεψης.
- ο Η χρήση ενός ενιαίου νευρωνικού δικτύου εκπαιδευμένου από όλες τις τιμές των χρονοσειρών έδωσε κατά μέσο όρο τα χειρότερα αποτελέσματα συγκριτικά με τον συνδυασμό στατιστικών μεθόδων και αυτόματης επιλογής ξεχωριστής στατιστικής μεθόδου. Αυτό το γεγονός μπορεί να οφείλεται στην έλλειψη επαρκούς όγκου δεδομένων.
- ο Ο συνδυασμός διαφορετικών παραμέτρων που διαμόρφωσε τα μοντέλα διαφορετικής αρχιτεκτονικής του νευρωνικού δικτύου είχαν μια μικρή επίδραση στη διαμόρφωση του τελικού σφάλματος.
- ο Συνολικά, από όλες τις υλοποιήσεις που πραγματοποιήθηκαν ο ισοβαρής συνδυασμός των προβλέψεων των μεθόδων sNaive, ETS , ARIMA παρουσίασε το πιο μικρό σφάλμα sMAPE= 13,261%

Προοπτικές και Μελλοντικές Προεκτάσεις

5.2 Στα πλαίσια μιας διπλωματικής εργασίας είναι αδύνατο, τόσο χρονικά όσο και λόγω της περιορισμένης πρόσβασης στα εταιρικά ευαίσθητα δεδομένα, να εξεταστούν όλες οι πτυχές του προβλήματος και να δοθεί μια ολοκληρωμένη λύση η οποία καλύπτει το σύνολο των επιμέρους θεμάτων στη πρόβλεψη ζήτησης για τον προγραμματισμό παραγωγής.

Κατά την υλοποίηση της εργασίας εντοπίστηκαν τα παρακάτω θέματα τα οποία θεωρούμε ότι χρήζουν περαιτέρω διερεύνησης και θα μπορέσουν να βελτιώσουν τα συνολικά αποτελέσματα.

- Αναζήτηση μεγαλύτερου εύρους καταγεγραμμένων δεδομένων μηνιαίας συχνότητας αλλά και περεταίρω εμβάθυνση σε ημερήσια δεδομένα. Στην παρούσα διπλωματική εργασία ήταν διαθέσιμα δεδομένα μόνο από το έτος 2017 έως και το 2019. Όσο μεγαλύτερο εύρος δεδομένων διαθέτουμε τόσο πιο εύκολα μπορούν να εντοπιστούν τα χαρακτηριστικά κάθε χρονοσειράς και να οδηγήσουν σε πιο ακριβείς προβλέψεις.
- Μελέτη επιπλέον μοντέλων μηχανικής μάθησης. Ο μεγαλύτερος όγκος δεδομένων των επιμέρους χρονοσειρών θα έχει ως αποτέλεσμα τη δυνατότητα αξιοποίησης περισσότερων αλγορίθμων μηχανικής μάθησης.
- Αναζήτηση σχέσης κωδικού με μηχανή για πρόβλεψη παραγωγικής δυναμικότητας.
- Ανάλυση σε επίπεδο κωδικού και κατηγοριοποίηση αναλόγως του κόστους(ABC).
- Πρόβλεψη πωλήσεων ανά πελάτη. Περαιτέρω εμβάθυνση σε στρατηγικά σημαντικούς πελάτες
- Κατανόηση της φύσης των brands και ευρύτερη μελέτη ειδικών γεγονότων, δράσεων και άλλων ποιοτικών χαρακτηριστικών για την κατανόηση του μοτίβου των διαθέσιμων χρονοσειρών
- Παραγωγή κριτικών προβλέψεων και συνδυασμός τους με τα αποτελέσματα κλασσικών μεθόδων πρόβλεψης λαμβάνοντας υπόψη το γενικότερο οικονομικό κλίμα και τις κοινωνικές συνθήκες που επικρατούν στην Ελλάδα.
- Συνδυασμός μεθόδων πρόβλεψης με χρήση συντελεστών βαρύτητας, ανάλογα με τη σχετική ακρίβεια της κάθε μεθόδου και τη συν διακύμανση των σφαλμάτων πρόβλεψης, για καλύτερη προσέγγιση της πραγματικότητας.

Βιβλιογραφία

- Ασημακόπουλος, Β., & Πετρόπουλος, Φ. (2013). *Επιχειρησιακές Προβλέψεις*. Αθήνα: Εκδόσεις Συμμετρία.
- Assimakopoulos, V., & Nikolopoulos, K. (2000). *The theta model: A decomposition approach to forecasting*. International Journal of Forecasting 16 (2000) 521–53.
- Blomgren, O., & Eriksson, S. (2016). *Improving Demand Planning and Inventory Management A Supply Chain Evaluation from a Distributor's Perspective*. Gothenburg, Sweden: Department of Technology Management and Economics, CHALMERS UNIVERSITY OF TECHNOLOGY .
- Chase, C. W. (2013). *Demand-Driven Forecasting: A Structured Approach to Forecasting*. Wiley.
- Hyndman, R. J., & Athanasopoulos, G. (2018 2nd edition). *Forecasting: principles and practice*. Monash University, Australia: OTEXTS.
- Hyndman, R. J., & Khandakar, Y. (2008). *Automatic Time Series Forecasting: the forecast Package for R*. Journal of Statistical Software.
- Hyndman, R. J., Koehler, A. B., Snyder, R. D., & Grose, S. (2002). *A state space framework for automatic forecasting using exponential smoothing methods*. International Journal of Forecasting.
- Hyndman, R., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O'Hara-Wild, M., . . . Zhou, Z. (2020). *Forecasting Functions for Time Series and Linear Models*.
- Jae, S. K., & Joel, S. G. (2011). *OPERATIONS MANAGEMENT*. ΚΛΕΙΔΑΡΙΘΜΟΣ.
- Kenworthy, J. (1998). *Planning and Control of Manufacturing Operations*. Wiley.
- Lawrence, K. D., & Klimberg, R. K. (2010). *Advances in Business and Management Forecasting*. Emerald Group Publishing Limited.
- Makridakis, S., Wheelwright, S., & Hyndman, R. (1998, 3rd ed.). *Forecasting: Methods and Applications*. New York: John Wiley and Sons.
- Mentzer, J. T., & Bienstock, C. (1998). *Sales Forecasting Management*. Sage Publications.
- Pekša, J. (2018). *Extensible Portfolio of Forecasting Methods for ERP Systems: An Integration Approach*. Information Technology and Management Science.

- Q. Cao, R. B. (2016). *Data Driven Production Forecasting Using Machine Learning*. SPE Argentina Exploration and Production of Unconventional Resources Symposium.
- Shumway, R. H., & Stoffer, D. S. (2011 3rd Edition). *Timeseries Analysis ant Its Applications*. Springer Science+Business Media.
- Simon, H. (2010, 3η Έκδοση). *Νευρωνικά Δίκτυα & Μηχανική Μάθηση*. Εκδόσεις Παπασωτηρίου.
- Steven Nahmias, T. L. (2015(7th edition)). *Production and Operations Analysis*. Waveland Press, Inc.
- Turbide, D. (1993(1st ed.)). *MRP+: The Adaptation, Enhancement, and Application of MRP II*. Industrial Press, Inc.
- Καββαδίας Κωνσταντίνος. (2019). *Πρόβλεψη ζήτησης και προγραμματισμός παραγωγής*.
- Κοκκίνου, Λ. Θ. (2013). *Τεχνητά Νευρωνικά Δίκτυα και Εφαρμογές στα Συστήματα Αυτόματου Ελέγχου*.
- Μήτση, Ε. (2020). *Αξιολόγηση Μεθόδων Πρόβλεψης της Τιμής Ηλεκτρικής Ενέργειας στην Αγορά της Ουγγαρίας*.
- Πάππης, Κ. (2006). *Προγραμματισμός Παραγωγής*. Σταμούλης.
- Σεργέδος, Κ. (2020). *Πρόβλεψη Ζήτησης σε καταστάματα λιανικής με μεθόδους ιεραρχικής συνάθροισης και τεχνικές μηχανικής μάθησης*.
- Φλαμπούρη, Β. (2020). *Βελτιστοποίηση της διαδικασίας διαχείρισης αποθεμάτων μέσω της χρήσης προηγμένων αλγορίθμων πρόβλεψης*.
- Ολοκληρωμένα Αυτοπαλινδρομικά Μοντέλα Κινητού Μέσου Όρου (ARIMA)
Μέθοδος Εκθετικής Εξομάλυνσης (Exponential Smoothing)
Μέθοδος Theta - Διαγωνισμοί Προβλέψεων
[Μονάδα Προβλέψεων και Στρατηγικής - Τεχνικές Προβλέψεων \(fsu.gr\)](http://fsu.gr)
 - A state space framework for automatic forecasting using exponential smoothing method [hksg.pdf \(robjhyndman.com\)](http://robjhyndman.com)
 - [Develop a Structured Consensus Demand-Planning Process to Drive Revenue and Profitability Attainment \(gcom.cloud\)](http://gcom.cloud)

- [ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ ΚΑΙ ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ_Data.pdf \(sev.org.gr\)](#)
- [Analytics & Τεχνητή Νοημοσύνη Παρατηρητήριο Ψηφιακού Μετασχηματισμού ΣΕΒ \(sev.org.gr\)](#)
- [Forecast Strategies - SAP Help Portal](#)
- [Demand forecasting overview - Supply Chain Management | Dynamics 365 | Microsoft Docs](#)
- [A forecast ensemble benchmark | Rob J Hyndman](#)
- R Documentation and manuals [www.rdocumentation.org](#)