

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΛΟΠΟΝΝΗΣΟΥ  
ΤΜΗΜΑ ΛΟΓΙΣΤΙΚΗΣ & ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ

## Πρόγραμμα Μεταπτυχιακών Σπουδών στη Λογιστική και Χρηματοοικονομική



### Μεταπτυχιακή Διατριβή

Η μηχανική μάθηση στον εντοπισμό και την πρόληψη της  
φοροδιαφυγής

**Μαστοράκη Παναγιώτα**

Επιβλέπων Καθηγητής: **Νικολόπουλος Σωτήριος**

Διατριβή υποβληθείσα στο Τμήμα Λογιστικής & Χρηματοοικονομικής του  
Πανεπιστημίου Πελοποννήσου. Η παρούσα διατριβή αποτελεί μέρος των  
απαιτήσεων για την απόκτηση του Μεταπτυχιακού Διπλώματος στη Λογιστική και  
Χρηματοοικονομική

Καλαμάτα, Απρίλιος 2024

**Πρόγραμμα Μεταπτυχιακών Σπουδών στη Λογιστική και  
Χρηματοοικονομική**



**ΠΜΣ στη Λογιστική και Χρηματοοικονομική**

ΤΜΗΜΑ ΛΟΓΙΣΤΙΚΗΣ ΚΑΙ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ • ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΛΟΠΟΝΝΗΣΟΥ

**Τριμελής Εξεταστική Επιτροπή**

**Νικολόπουλος Σωτήριος (Επιβλέπων)**

**Επίκουρος Καθηγητής, Τμήμα Λογιστικής και Χρηματοοικονομικής,  
Πανεπιστήμιο Πελοποννήσου**

**Σηλιόπουλος Οδυσσέας**

**Αναπληρωτής Καθηγητής, Τμήμα Λογιστικής και Χρηματοοικονομικής,  
Πανεπιστήμιο Πελοποννήσου**

**Συρμαλόγλου Αδαμάντιος**

**Εργαστηριακό Διδακτικό Προσωπικό, Τμήμα Λογιστικής και  
Χρηματοοικονομικής, Πανεπιστήμιο Πελοποννήσου**

UNIVERSITY OF PELOPONNESE  
DEPARTMENT OF ACCOUNTING & FINANCE

**Master Program in Accounting and Finance**



**Master of Science (M.Sc)  
in Accounting and Finance**

**Postgraduate Thesis**

**Machine learning in detecting and preventing tax evasion**

**Mastoraki Panagiota**

**Supervising Professor: Nikolopoulos Sotirios**

This thesis submitted to the Department of Accounting & Finance of the University of Peloponnese.  
This dissertation is part of the requirements for obtaining the Master's Degree in Accounting and  
Finance

Kalamata, April 2024

UNIVERSITY OF PELOPONNESE  
DEPARTMENT OF ACCOUNTING & FINANCE

**Master Program in Accounting and Finance**



**Master of Science (M.Sc)  
in Accounting and Finance**

**Thesis Committee**

**Nikolopoulos Sotirios**

**Assistant Professor, Department of Accounting and Finance, University of  
Peloponnese**

**Spiliopoulos Odysseas**

**Associate Professor, Department of Accounting and Finance, University of  
Peloponnese**

**Syrmaloglou Adamantios**

**Laboratory Teaching Staff, Department of Accounting & Finance, University of  
Peloponnese**

δηλώνω υπεύθυνα ότι:

- 1)** Είμαι ο κάτοχος των πνευματικών δικαιωμάτων της πρωτότυπης αυτής εργασίας και από όσο γνωρίζω η εργασία μου δε συκοφαντεί πρόσωπα, ούτε προσβάλλει τα πνευματικά δικαιώματα τρίτων.
  
- 2)** Αποδέχομαι ότι το Τμήμα Λογιστικής & Χρηματοοικονομικής μπορεί, χωρίς να αλλάξει το περιεχόμενο της εργασίας μου, να τη διαθέσει σε ηλεκτρονική μορφή μέσα από τη ψηφιακή Βιβλιοθήκη του Ιδρύματος, να την αντιγράψει σε οποιοδήποτε μέσο ή/και σε οποιοδήποτε μορφότυπο καθώς και να κρατά περισσότερα από ένα αντίγραφα για λόγους συντήρησης και ασφάλειας.

## Πίνακας περιεχομένων

Περίληψη στα Ελληνικά .....	vii
Περίληψη στα Αγγλικά .....	viii
Εισαγωγή .....	1
<b>Κεφάλαιο 1<sup>ο</sup>: Μηχανική Μάθηση .....</b>	<b>3</b>
1.1 Ιστορική Αναδρομή Μηχανικής Μάθησης.....	3
1.2 Ορισμός Μηχανικής Μάθησης .....	4
1.3 Είδη Μηχανικής Μάθησης .....	6
1.4 Αλγόριθμοι Μηχανικής Μάθησης .....	8
1.5 Πλεονεκτήματα και μειονεκτήματα της μηχανικής μάθησης.....	13
1.6 Μηχανική Μάθηση ως Αντικείμενο Μελέτης.....	15
<b>Κεφάλαιο 2<sup>ο</sup>: Φορολογία και Φοροδιαφυγή .....</b>	<b>16</b>
2.1 Ιστορική εξέλιξη φορολογίας.....	16
2.2 Φορολογία.....	19
2.2.1 Βασικές έννοιες.....	19
2.2.2 Βασικές λειτουργίες των φόρων.....	20
2.2.3 Κανόνες ενός φορολογικού συστήματος .....	21
2.2.4 Σκοπός της φορολογίας.....	22
2.3 Φοροδιαφυγή .....	23
2.3.1 Ορισμός Φοροδιαφυγής .....	23
2.3.2 Φοροδιαφυγή και Φοροαποφυγή.....	23
2.3.3 Φοροδιαφυγή και Παραοικονομία.....	24
2.3.4 Φοροδιαφυγή και Εγκληματικές Δραστηριότητες .....	25
2.4 Τρόποι μείωσης φοροδιαφυγής.....	25
<b>Κεφάλαιο 3<sup>ο</sup>: Μηχανική Μάθηση και Φοροδιαφυγή .....</b>	<b>29</b>
3.1 Πρόσφατη αρθρογραφία.....	29

3.2	Ο ρόλος των τεχνικών μηχανικής μάθησης στον εντοπισμό και την πρόληψη της φοροδιαφυγής.....	33
3.3	Μηχανική μάθηση: εφαρμογή στη λογιστική διαχείριση.....	36
3.4	Υπάρχουσες εφαρμογές εντοπισμού και πρόληψης της φοροδιαφυγής .....	38
<b>Κεφάλαιο 4<sup>ο</sup>: Μελέτες περίπτωσης μηχανικής μάθησης στη φοροδιαφυγή .....</b>		<b>41</b>
4.1	Ελλάδα.....	41
4.2	Μεξικό .....	44
4.3	Ισπανία.....	45
4.4	Ινδία.....	46
4.5	Τουρκία.....	47
4.6	Πολωνία.....	47
<b>Συμπεράσματα και Προτάσεις.....</b>		<b>49</b>
<b>Βιβλιογραφία .....</b>		<b>51</b>

## Περίληψη στα Ελληνικά

Στη σύγχρονη εποχή, που τα πάντα εξελίσσονται με ταχύτατους ρυθμούς, ήταν σχεδόν αναμενόμενο να αναπτυχθεί και να διαδοθεί η τεχνική νοημοσύνη με ιδιαίτερη έμφαση στη μηχανική μάθηση. Ωστόσο, αυτή η διάδοση που οι υπολογιστές μαθαίνουν από τα δεδομένα και βελτιώνονται με την εμπειρία χωρίς να είναι σαφώς προγραμματισμένοι, έχει αρχίσει να συνδέεται όλο και περισσότερο και με άλλες επιστήμες. Η φοροδιαφυγή αποτελεί ένα από τα σημαντικά προβλήματα που αντιμετωπίζουν οι φορολογικές αρχές παγκοσμίως, έχοντας αρνητική επίδραση στις δημοσιονομικές πολιτικές. Η εφαρμογή της μηχανικής μάθησης στον τομέα της φορολογίας φαίνεται να είναι καθοριστική για την αντιμετώπιση αυτού του προβλήματος. Στη συγκεκριμένη εργασία γίνεται προσπάθεια σύνδεσης της μηχανικής μάθησης με τη φοροδιαφυγή.

**Λέξεις κλειδιά:** Μηχανική μάθηση, Φορολογία, Φοροδιαφυγή



## Περίληψη στα Αγγλικά

In today's fast-paced world, where everything seems to be evolving at breakneck speed, the development and spread of artificial intelligence, particularly machine learning, was almost a foregone conclusion. This ability of computers to learn from data and improve on their own, without needing explicit programming, has opened doors to its use in many different fields. Tax evasion is a major problem for tax authorities around the globe, as it hurts government finances. Machine learning applied to taxation seems to be a game-changer in the fight against this issue. This specific study aims to delve into the connection between machine learning and tax evasion.

**Keywords:** Machine learning, Taxation, Tax evasion

## Εισαγωγή

Στις μέρες μας, οι έννοιες της μηχανικής μάθησης και της φοροδιαφυγής είναι ευρέως διαδεδομένες και η χρήση τους βοηθά τους εφοριακούς και τους λογιστές στις καθημερινές τους πρακτικές. Η φοροδιαφυγή είναι ένα πολύ σοβαρό οικονομικό έγκλημα και η μηχανική μάθηση μπορεί να αξιοποιηθεί για τον εντοπισμό και την πρόληψη αυτών των περιπτώσεων.

Η εφαρμογή της μηχανικής μάθησης στον τομέα της φορολογίας μπορεί να συμβάλει σημαντικά στον εντοπισμό και την αντιμετώπιση περιπτώσεων φοροδιαφυγής, ενισχύοντας μια πιο διαφανή και δίκαιη προσέγγιση για την επιβολή του φόρου. Με τη χρήση μηχανικής μάθησης, οι φορολογικές αρχές μπορούν να αποτρέψουν πιθανές παραβιάσεις και να διενεργήσουν εντατικούς ελέγχους σε συγκεκριμένους φορολογούμενους ή εταιρείες που παρουσιάζουν ύποπτη συμπεριφορά. Ακόμα, είναι δυνατή η ανάλυση μεγάλων συνόλων δεδομένων φορολογικών πληροφοριών για τον εντοπισμό ανωμαλιών και μοτίβων που μπορεί να υποδηλώνουν περιπτώσεις φοροδιαφυγής. Αυτή η δυνατότητα επιτρέπει στις φορολογικές αρχές να εστιάσουν τους πόρους τους σε πιο στρατηγικούς και αποτελεσματικούς ελέγχους, βοηθώντας έτσι στην εξασφάλιση της τήρησης της φορολογικής νομοθεσίας.

Η φοροδιαφυγή είναι ένα σοβαρό ζήτημα σε πολλές χώρες. Αυτό οδηγεί σε μια πιο αποτελεσματική και δικαιότερη φορολογική επιβολή, ενισχύοντας τη χρηματοοικονομική διαφάνεια και διατηρώντας την οικονομική ισορροπία των κρατών. Σε ότι αφορά την Ελλάδα, διάφορες πτυχές της εφαρμογής τεχνικών μηχανικής μάθησης, είναι ο εντοπισμός προτύπων φοροδιαφυγής, η ανάλυση οικονομικών δεδομένων σε πραγματικό χρόνο, η παρακολούθηση των φορολογουμένων, και η επεξεργασία φορολογικών δηλώσεων.

Η σύνδεση της μηχανικής μάθησης με τη φοροδιαφυγή μπορεί να γίνει με διαφορετικούς τρόπους, ανάλογα με τη φύση του προβλήματος και τον αλγόριθμο που χρησιμοποιείται. Η συχνότητα και οι προκλήσεις που εγκυμονούν στο φαινόμενο της φοροδιαφυγής, καθώς και η ανάγκη για άμεσες και αποτελεσματικές λύσεις ενισχύουν τη χρήση της μηχανικής μάθησης στον τομέα της φορολογίας. Σκοπός και αντικείμενο της παρούσας μελέτης είναι η διερεύνηση της μηχανικής μάθησης στον εντοπισμό και την πρόληψη της φοροδιαφυγής.

Η εργασία θα ολοκληρωθεί μέσα από μελέτη αρθρογραφίας και βιβλιογραφίας με σκοπό να προσφέρει νέες γνώσεις πάνω στο θέμα σύμφωνα με πρόσφατες μελέτες.

Πιο αναλυτικά και σύμφωνα με τη διάρθρωση της εργασίας, θα ολοκληρωθεί μέσα από πέντε κεφάλαια.

Αρχικά, στο **πρώτο κεφάλαιο** γίνεται αναφορά στη μηχανική μάθηση ως προς την ιστορική της εξέλιξη, ορισμοί που έχουν δοθεί κατά καιρούς, στα είδη της όπως και στους αλγόριθμους. Ακόμα, αναφέρονται τα πλεονεκτήματα και τα μειονεκτήματα που έχει ως κλάδος.

Στη συνέχεια, στο **δεύτερο κεφάλαιο** παρέχει μια γενική εισαγωγή στο θέμα της φορολογίας με ιδιαίτερη αναφορά σε θέματα φοροδιαφυγής, όπως τον ορισμό αυτής, την ιστορία, το κόστος που αυτή συνεπάγεται και τον διαχωρισμό αυτής της έννοιας με άλλες που συνήθως γίνεται σύγχυση.

Το **τρίτο κεφάλαιο** παρουσιάζεται η κυριότερη πρόσφατη αρθρογραφία σχετικά με την σύνδεση των δύο εννοιών και έπειτα αναλύεται ο ρόλος των τεχνικών μηχανικής μάθησης στον εντοπισμό και την πρόληψη της φοροδιαφυγής όπως αυτή προκύπτει από την ανάλυση της διαθέσιμης βιβλιογραφίας. Τέλος, θα αναφερθούν εφαρμογές της μηχανικής μάθησης στη λογιστική διαχείριση αλλά και στη φοροδιαφυγή.

Στο **τέταρτο κεφάλαιο** γίνεται αναφορά σε συγκεκριμένες εφαρμογές με παραδείγματα στην Ελλάδα, το Μεξικό, την Ισπανία, την Ινδία, την Τουρκία και την Πολωνία.

Η μελέτη θα κλείσει καταλήγοντας στα **συμπεράσματα** και στις προτάσεις.

# Κεφάλαιο 1<sup>ο</sup>:Μηχανική Μάθηση

Η μηχανική μάθηση είναι ένα υποπεδίο της τεχνητής νοημοσύνης, το οποίο ορίζεται ευρέως ως η ικανότητα μιας μηχανής να μιμείται την ευφυή ανθρώπινη συμπεριφορά. Σκοπός της είναι η ανίχνευση μοτίβων και να μαθαίνει πώς να κάνει προβλέψεις και συστάσεις μέσω της επεξεργασίας δεδομένων. Είναι αναγκαίο να διατυπωθούν δύο διαπιστώσεις για την μηχανική μάθηση: ο διεπιστημονικός της χαρακτήρας που αντλείται από κάθε είδους επιστήμης πληροφορίες, και ο κίνδυνος του ερευνητή να επιμείνει σε μια δεδομένη οπτική του αλγόριθμου πολύ έντονα, αδιαφορώντας για τις υπόλοιπες.

## 1.1 Ιστορική Αναδρομή Μηχανικής Μάθησης

Η μηχανική μάθηση αναπτύχθηκε στα μέσα του 20ου αιώνα, όταν οι ερευνητές άρχισαν να διερευνούν πώς να κάνουν τις μηχανές να μαθαίνουν από δεδομένα. Αφορά την εκμάθηση από δεδομένα και τη λήψη αποφάσεων ή προβλέψεων. Ουσιαστικά βασίζεται στην υπόθεση ότι οι μηχανές μπορούν να αποκτήσουν ευφυΐα που τους επιτρέπει να μάθουν από προηγούμενους υπολογισμούς και να προσαρμοστούν στο περιβάλλον (Ανδρώνης, 2021).

Η ιστορία της μηχανικής μάθησης χρονολογείται από το 1943 όταν ο McCulloch πρότεινε ένα μαθηματικό μοντέλο από στο νευρωνικό δίκτυο (Neural Network) για υπολογιστές (Nayak and Patgiri, 2021). Ωστόσο, ο όρος "μηχανική μάθηση" χρησιμοποιήθηκε για πρώτη φορά από τον Arthur Samuel το 1959, ο οποίος την όρισε ως την ικανότητα μιας μηχανής να μαθαίνει χωρίς να προγραμματίζεται ρητά (Helm et al., 2020). Τις επόμενες δεκαετίες αναπτύχθηκαν διάφορες προσεγγίσεις στη μηχανική μάθηση, όπως τα δέντρα αποφάσεων και τα δίκτυα Bayes. Ωστόσο, η πρόοδος στη μηχανική μάθηση ήταν αργή λόγω των περιορισμών στην υπολογιστική ισχύ και την αποθήκευση δεδομένων (Dong et al., 2021).

Στη δεκαετία του 1990, με την έλευση ισχυρότερων υπολογιστών και τη διαθεσιμότητα μεγάλων ποσοτήτων δεδομένων, η μηχανική μάθηση άρχισε να κάνει σημαντικά βήματα προόδου. Εισήχθησαν οι μηχανές διανυσμάτων υποστήριξης (SVMs), που είναι μία ομάδα αλγορίθμων μάθησης με επίβλεψη που αρχικά χρησιμοποιήθηκαν για την κατηγοριοποίηση ενώ αργότερα εφαρμόστηκαν και σε προβλήματα παλινδρόμησης, και οι μέθοδοι συνόλων, όπως τα τυχαία δάση (Παπαποστόλου, 2017). Ακόμα γεννήθηκε ο τομέας της βαθιάς μάθησης με την ανάπτυξη των νευρωνικών δικτύων.

Στις αρχές της δεκαετίας του 2000, η αύξηση των μεγάλων δεδομένων και η διαθεσιμότητα του υπολογιστικού νέφους (cloud computing), δηλαδή η διάθεση υπολογιστικών πόρων μέσω διαδικτύου, επιτάχυναν περαιτέρω την ανάπτυξη της μηχανικής μάθησης. Αναπτύχθηκαν νέες τεχνικές όπως η ενισχυτική μάθηση, τα δημιουργικά αντιφατικά δίκτυα (generative adversarial) και η μάθηση μεταφοράς (transfer learning), επιτρέποντας στις μηχανές να μαθαίνουν πιο αποδοτικά και αποτελεσματικά από μεγάλες ποσότητες δεδομένων (Rathore et al., 2021). Σήμερα, η μηχανική μάθηση χρησιμοποιείται σε ένα ευρύ φάσμα εφαρμογών, από την αναγνώριση εικόνας και ομιλίας έως την επεξεργασία φυσικής γλώσσας και τα αυτόνομα οχήματα. Καθώς η υπολογιστική ισχύς και η αποθήκευση δεδομένων συνεχίζουν να αυξάνονται και επιπλέον επειδή αναπτύσσονται νέες τεχνικές, είναι πιθανό η μηχανική μάθηση να έχει μεγαλύτερη επίδραση τα επόμενα χρόνια.

## 1.2 Ορισμός Μηχανικής Μάθησης

Όταν εμφανίστηκαν οι υπολογιστές, οι ερευνητές άρχισαν να μελετούν και να προσπαθούν να εμφυτεύσουν δυνατότητες μάθησης σε υπολογιστές (Carbonell, Michalski and Mitchell, 1983). Ένα από τα χαρακτηριστικά της μηχανικής μάθησης είναι ότι βρίσκεται στα όρια πολλών διαφορετικών κλάδων, κυρίως της επιστήμης των υπολογιστών, της στατιστικής, των μαθηματικών και της μηχανικής (Herbrich and Graepel, 2015). Ως διαδικασία, αφορά στην απόκτηση νέων ή στην τροποποίηση ήδη υπάρχοντων συμπεριφορών, αξιών, δεξιοτήτων και προτιμήσεων. Μπορεί να σημειωθεί πως η εφαρμογή της μηχανικής μάθησης έχει ενταχθεί στην καθημερινότητα τα τελευταία χρόνια, όπως για παράδειγμα μέσα από αυτόματες προτάσεις για ταινίες, προτάσεις παραγγελιών φαγητού, αναγνώριση προσώπων σε ψηφιακές φωτογραφίες (Κούτσης, 2022).

Διάφοροι ερευνητές κατά διαστήματα έχουν προσεγγίσει την έννοια αυτή, δίνοντας διαφορετικούς ορισμούς και τονίζοντας κάθε φορά μερικές χρήσιμες και ποικίλες προοπτικές του τομέα αυτού. Ενδεικτικά θα γίνει αναφορά σε ορισμένους από αυτούς. Από τους πρώτους ορισμούς που δίνονται για τη μηχανική μάθηση είναι του Arthur Samuel το 1959 που αναφέρει ότι η μηχανική μάθηση είναι «Το πεδίο σπουδών που επιτρέπει τους υπολογιστές να μαθαίνουν χωρίς ρητό προγραμματισμό» (Alzubi, Nayyar and Kumar, 2018). Αυτή η δήλωση συμπληρώθηκε από τον Tom Mitchell, ο οποίος περιγράφει τη διαδικασία μηχανικής μάθησης με ένα πιο αυστηρό μαθηματικό τρόπο:

*«What a machine does is, it takes a **task** (T), executes it, and **measures its performance** (P). Now a machine has a large number of data, so as it processes that data, its **experience** (E)*

*increases over time, resulting in a higher performance measure (P). So after going through all the data, our machine learning model's accuracy increases, which means that the predictions made by our model will be very accurate. »*

Ο παραπάνω ορισμός ορίζει την τυπική διαδικασία της μηχανικής μάθησης, δηλαδή να προσέχουμε τα δεδομένα που θα εισάγουμε να είναι σωστά, ποιες αποφάσεις πρέπει να λάβει το λογισμικό και πώς θα αξιολογήσουμε τα αποτελέσματά του (Brownlee, 2013). Ακόμα, αναφέρεται ότι είναι «Ο τομέας της μηχανικής μάθησης ασχολείται με το ζήτημα του τρόπου κατασκευής προγραμμάτων υπολογιστών που βελτιώνονται αυτόματα με την εμπειρία.» (Mitchell, 1997).

Ο Alpaydin αναφέρει πως η μηχανική μάθηση θεωρείται ο προγραμματισμός ενός υπολογιστή με σκοπό τη βελτίωση κάποιου κριτηρίου απόδοσης χρησιμοποιώντας ως δεδομένα παρελθοντικές εμπειρίες (Alpaydin, 2010). Το μοντέλο έχει φτιαχτεί από αρχής με διάφορες παραμέτρους, τις οποίες αλλάζει και βελτιώνει μέσω της εμπειρίας και των νέων δεδομένων που διαχειρίζεται. Ενδιαφέρεται για τη «βελτίωση της γνώσης και της απόδοσης των μηχανών μέσω της αλληλεπίδρασης με το περιβάλλον, και τα δεδομένα», με ελάχιστη ή καθόλου ανθρώπινη παρέμβαση (Fossaceca and Young, 2018). Χρησιμοποιεί αλγόριθμους που μαθαίνουν επαναληπτικά από δεδομένα προκειμένου να βελτιώσουν τις προβλέψεις τους. Καθώς οι αλγόριθμοι χρησιμοποιούν μεγάλο όγκο δεδομένων είναι πλέον σε θέση να παράγουν πιο ακριβείς προβλέψεις με βάση αυτά τα δεδομένα, τα οποία ονομάζονται «δεδομένα εκπαίδευσης» (Poole, Mackworth and Goebel, 1998). Τα συστήματα δεν είναι από την αρχή ικανά να χρησιμοποιήσουν από μόνα τους δεδομένα. Χρειάζεται μια πρώτη εκπαίδευση κατά την οποία το σύστημα τροφοδοτείται με διάφορα δεδομένα και έτσι μαθαίνει να συνδυάζει και να βρίσκει μοτίβα στα δεδομένα που θα το βοηθήσουν σε καλύτερες επιλογές και αποφάσεις στο μέλλον (Alpaydin, 2010).

Μια ακόμα ερμηνεία της μηχανικής μάθησης δίνεται από τους Herbrich και Graepel που αναφέρουν ότι «Η μηχανική μάθηση μελετάται συνήθως ως μέρος της τεχνητής νοημοσύνης, η οποία την τοποθετεί σταθερά στην επιστήμη των υπολογιστών (Herbrich and Graepel, 2015).

Η υπολογιστική στατιστική είναι ο δεσμός μεταξύ της στατιστικής και της επιστήμης των υπολογιστών και χρησιμοποιούν στις στατιστικές μεθόδους που ενεργοποιούνται με τη χρήση υπολογιστικών μεθόδων. Συνδέεται επίσης με τη Μαθηματική Βελτιστοποίηση (Mathematical Optimization), η οποία ενσωματώνει μοντέλα, εφαρμογές και τεχνολογικό πλαίσιο στον τομέα της στατιστικής (Alzubi, Nayyar and Kumar, 2018). Έτσι, ο κύριος ρόλος του στατιστικολόγου

είναι η ερμηνεία των δεδομένων, η παρουσίαση και η ανάλυση τους, εφαρμόζοντας στατιστικές μεθόδους και στατιστικά εργαλεία με στόχο την εξαγωγή συμπερασμάτων και προβλέψεων, απαραίτητων για τη λήψη ορθών αποφάσεων (Παπακωνσταντίνου and Καϊτσα, 1997).

### 1.3 Είδη Μηχανικής Μάθησης

Η μηχανική μάθηση αφορά την ικανότητα δημιουργίας αλγορίθμων που μπορούν να μάθουν και να κάνουν προβλέψεις με βάση τα δεδομένα που λαμβάνουν. Αξίζει να σημειωθεί ότι διακρίνονται σε μάθηση με επίβλεψη (Supervised), μάθηση χωρίς επίβλεψη (Unsupervised), μάθηση με ημι-επίβλεψη (Semi-Supervised), Ενισχυτική μάθηση και βαθιά μάθηση (Γεωργούλη, 2015), τα οποία αναλύονται συνοπτικά παρακάτω:

**«Μάθηση με επίβλεψη» (Supervised):** Η εποπτευόμενη μάθηση χρησιμοποιείται συνήθως σε σενάρια με αρκετά επισημασμένα δεδομένα. Είναι χρήσιμη όταν εκπαιδεύεται ένα μοντέλο ώστε να παράγει λογικές προβλέψεις ως απάντηση σε νέα δεδομένα (Bishop, 2006). Περιλαμβάνει έναν αλγόριθμο που χρησιμοποιεί ένα σύνολο δεδομένων εκπαίδευσης και λαμβάνει ανατροφοδότηση από ανθρώπους για να μάθει τη σχέση δεδομένων εισροών με μια δεδομένη έξοδο. Άρα, θα μπορούσαμε να πούμε ότι οι μηχανές λαμβάνουν επισημασμένες εισόδους και τις επιθυμητές εξόδους τους. Ο στόχος είναι να μάθουν έναν γενικό κανόνα για την αντιστοίχιση των εισόδων στην έξοδο. Η μέθοδος έχει χρησιμοποιηθεί ευρέως σε πολλά πεδία, όπως αναγνώριση αντικειμένων, αναγνώριση ομιλίας και ανίχνευση ανεπιθύμητων μηνυμάτων. Τυπικοί αλγόριθμοι σε αυτήν την κατηγορία περιλαμβάνουν δέντρα αποφάσεων (Decision Tree), K-πλησιέστερου γείτονα (K-Nearest Neighbor), τυχαίου δάσους (Random Forest), Naive Bayes και νευρωνικά δίκτυα (Neural Network), που θα αναλυθούν σε επόμενη υποενότητα (Kato et al., 2020).

**«Μάθηση χωρίς επίβλεψη» (Unsupervised):** Στους αλγορίθμους μάθησης χωρίς επίβλεψη παρέχεται ένα σύνολο εισόδου χωρίς ετικέτα για να συμπεράνει σωστά την έξοδο. Οι τεχνικές χρησιμοποιούνται συνήθως για ομαδοποίηση και συγκέντρωση. Σε μια προσέγγιση «χωρίς επίβλεψη μάθησης», ο αλγόριθμος θα διερευνήσει τα δεδομένα εισόδου χωρίς να του δοθεί μια ρητή μεταβλητή απόκρισης, όπως για παράδειγμα διερευνά δημογραφικά δεδομένα πελατών για να εντοπίσει μοτίβα. Σε αυτή την περίπτωση, ο υπεύθυνος λήψης αποφάσεων θέλει ο αλγόριθμος να βρει μοτίβα και να ταξινομήσει τους πελάτες. Είναι χρήσιμη όταν θέλουμε να εξερευνήσουμε τα δεδομένα, αλλά δεν έχουμε ακόμη συγκεκριμένο στόχο ή δεν είμαστε σίγουροι για το πληροφοριακό περιεχόμενο των δεδομένων. Τυπικοί μη

εποπτευόμενοι αλγόριθμοι περιλαμβάνουν το K-means, τους αυτο-οργανωμένους χάρτες (Self-Organizing Maps), το κρυφό μοντέλο Markov (Hidden Markov Model), η περιορισμένη μηχανή Boltzmann (Restricted Boltzmann machine) (Ali et al., 2020).

**«Μάθηση με ημι-επίβλεψη» (Semi-Supervised):** Αναφέρεται στην ιδέα της χρήσης ενός μεγάλου συνόλου μη επισημασμένων δεδομένων  $U$  για την ενίσχυση ενός δεδομένου συνόλου επισημασμένων δεδομένων  $L$ , προκειμένου να παραχθούν ακριβέστεροι κανόνες από ό,τι θα είχε επιτευχθεί με τη χρήση μόνο του  $L$  (Blum, Horcroft and Kannan, 2018). Ως τυπικό παράδειγμα του ημι-εποπτευόμενου αλγορίθμου μάθησης, η ψευδοσήμανση είναι απλή και αποτελεσματική. Τα περιορισμένα δεδομένα με σήμανση χρησιμοποιούνται για την εκπαίδευση ενός μοντέλου. Στη συνέχεια, το εκπαιδευμένο μοντέλο χρησιμοποιείται για τη δημιουργία ψευδών σημάνσεων για τα δεδομένα χωρίς σήμανση. Τέλος, τα επισημασμένα δεδομένα και τα ψευδο-επισημασμένα δεδομένα χρησιμοποιούνται για την επανεκπαίδευση του μοντέλου. Για να αξιοποιηθούν πλήρως τα δεδομένα χωρίς σήμανση, οι ημι-εποπτευόμενοι αλγόριθμοι μάθησης συνήθως απαιτούν ορισμένες παραδοχές στο σύνολο δεδομένων, όπως πολλαπλή υπόθεση, υπόθεση χαμηλής πυκνότητας, υπόθεση συμπλέγματος και υπόθεση ομαλότητας.

**«Ενισχυτική μάθηση» (Reinforcement algorithms, Q-learning, policy gradient methods, and actor-critic methods):** Αυτή η προσέγγιση είναι μια τεχνική μηχανικής μάθησης που επικεντρώνεται στην εκπαίδευση ενός αλγορίθμου ακολουθώντας την προσέγγιση «κόψε και δοκίμασε». Η κύρια ιδέα της μάθησης ενίσχυσης είναι να μιμηθεί τη μαθησιακή διαδικασία του εγκεφάλου με δοκιμή και σφάλμα. Αντί να μάθει τη δομή του συνόλου δεδομένων κατάρτισης, η μάθηση ενίσχυσης προσπαθεί να διερευνήσει τις καλύτερες ενέργειες κατά τη διάρκεια μιας δυναμικής διαδικασίας. Ασχολείται με το πρόβλημα της εύρεσης κατάλληλων ενεργειών που πρέπει να γίνουν σε μια δεδομένη κατάσταση προκειμένου να μεγιστοποιηθεί μια ανταμοιβή. Ουσιαστικά, υπάρχει μια ακολουθία καταστάσεων και ενεργειών κατά τις οποίες ο αλγόριθμος μάθησης αλληλοεπιδρά με το περιβάλλον του (Bishop, 2006). Η τεχνική αντιμετωπίζει ορισμένες προκλήσεις στο χειρισμό προβλημάτων σε ένα μεγάλο χώρο κατάστασης ή χώρο δράσης, επειδή είναι πολύ δύσκολο να μοντελοποιηθεί κάθε ζεύγος κατάστασης-δράσης άμεσα. Κατά συνέπεια, η μάθηση ενίσχυσης σπάνια χρησιμοποιείται στην πράξη.

**Βαθιά μάθηση:** Είναι ένας τύπος μηχανικής μάθησης που μπορεί να επεξεργαστεί ένα ευρύτερο φάσμα δεδομένων, απαιτεί λιγότερη προεπεξεργασία δεδομένων από τον άνθρωπο



και μπορεί συχνά να παράγει πιο ακριβή αποτελέσματα από τις παραδοσιακές προσεγγίσεις μηχανικής μάθησης (Dean, 2014). Η βαθιά μάθηση είναι μια συγκεκριμένη μέθοδος μηχανικής μάθησης που ενσωματώνει νευρωνικά δίκτυα σε διαδοχικά στρώματα προκειμένου να μαθαίνει από τα δεδομένα με επαναληπτικό τρόπο. Τα διασυνδεδεμένα στρώματα υπολογιστών που βασίζονται σε λογισμικό τα οποία ονομάζονται «νευρώνες», σχηματίζουν ένα νευρωνικό δίκτυο. Η βαθιά μάθηση διαχωρίζεται σε «συμβατικό νευρωνικό δίκτυο» και «επαναλαμβανόμενο νευρωνικό δίκτυο». Έτσι, οι μηχανές μαθαίνουν μέσω δοκιμής και σφάλματος χρησιμοποιώντας δεδομένα από τη δική τους εμπειρία.

#### 1.4 Αλγόριθμοι Μηχανικής Μάθησης

Στις μέρες μας, οι αλγόριθμοι μηχανικής μάθησης χρησιμοποιούνται όλο και περισσότερο. Πέρα των εμπορικών εφαρμογών, η μηχανική μάθηση αποτελεί ένα ισχυρό εργαλείο διάφορων επιστημονικών προβλημάτων, όπως είναι η πολυπλοκότητα του φορολογικού κώδικα.

Οι αλγόριθμοι προσαρμόζονται ως απάντηση σε νέα δεδομένα και εμπειρίες για να βελτιώσουν την αποτελεσματικότητα με την πάροδο του χρόνου (Bishop, 2006). Σε πολλές περιπτώσεις απαιτείται συνδυασμός αλγορίθμων μηχανικής μάθησης, που συνεπάγεται ταξινόμηση της αποκτηθείσας γνώσης.

Τα μοντέλα ταξινόμησης αφορούν τη δημιουργία μοντέλων που μπορούν να ταξινομήσουν δεδομένα που δεν έχουν ακόμη ταξινομηθεί. Δηλαδή, ο στόχος είναι να βρεθούν οι περιορισμοί απόφασης που καθορίζουν τον τρόπο διαχωρισμού των κλάσεων. Επίσης, μπορούν να χρησιμοποιηθούν ως ερμηνευτικά εργαλεία για τη διάκριση αντικειμένων που ανήκουν σε διαφορετικές κατηγορίες ή ως εργαλεία για την πρόβλεψη της κατηγορίας νέων εγγραφών (Tan *et al.*, 2018).

Σύμφωνα με τον Langley (1996), η οργάνωση της γνώσης μπορεί να κατηγοριοποιηθεί ως μια λίστα αποφάσεων. Στη συνέχεια, υπάρχουν δίκτυα συμπερασμάτων (inference networks), όπου τα δεδομένα είναι οργανωμένα σε ένα δέντρο, που σημαίνει ότι το εύρος κάθε κόμβου επηρεάζεται από τις συνδέσεις του με τους κόμβους κάτω από αυτόν. Τέλος, η οργάνωση δεδομένων μπορεί να γίνει με ιεραρχία ιδεών που αποτυπώνονται με τη μορφή δέντρων. Η διαφορά από την οργάνωση του δικτύου συμπερασμάτων είναι ότι εδώ κάθε κόμβος αντιστοιχεί σε μια ιδέα και στη σχετική περιγραφή της. Οι κόμβοι με γενικές περιγραφές κατατάσσονται υψηλότερα στην ιεραρχία, ενώ οι κόμβοι με πιο συγκεκριμένες περιγραφές κατατάσσονται χαμηλότερα (Langley, 1996).

## **01. Λογιστική παλινδρόμηση**

Η λογιστική παλινδρόμηση είναι ένας εποπτευόμενος αλγόριθμος μηχανικής μάθησης που ολοκληρώνει εργασίες δυαδικής ταξινόμησης προβλέποντας την πιθανότητα ενός αποτελέσματος, ενός συμβάντος ή μιας παρατήρησης. Η λογιστική παλινδρόμηση αναλύει τη σχέση μεταξύ μιας ή περισσότερων ανεξάρτητων μεταβλητών και ταξινομεί τα δεδομένα σε διακριτές κλάσεις. Χρησιμοποιείται συνήθως σε προβλήματα δυαδικής ταξινόμησης όπου η μεταβλητή αποτελέσματος αποκαλύπτει μία από τις δύο κατηγορίες (0 και 1) (van den Goorbergh *et al.*, 2022).

## **02. Δέντρο απόφασης**

Τα δέντρα αποφάσεων είναι μια από τις πιο γνωστές μεθόδους ταξινόμησης, η οποία αναπαριστάται από δομή δέντρου. Είναι εύκολο να κατανοηθούν και μπορούν να χρησιμοποιηθούν σε πολλές διαφορετικές εφαρμογές για την ανάλυση δεδομένων βάσει πιθανοτήτων για τη λήψη καλύτερων επιχειρηματικών αποφάσεων. Είναι επίσης εξαιρετικά ευέλικτα - μπορούν να χρησιμοποιηθούν για τα πάντα, από μάρκετινγκ έως ιατρικές διαγνώσεις. Ένα δέντρο αποφάσεων είναι μια γραφική αναπαράσταση πιθανών λύσεων σε μια απόφαση που βασίζεται σε ορισμένες συνθήκες. Υπάρχουν διάφοροι τύποι δέντρων αποφάσεων, που χρησιμοποιούνται τόσο για προβλήματα παλινδρόμησης όσο και για προβλήματα ταξινόμησης (Jijo and Abdulazeez, 2021).

Τα εποπτευόμενα δέντρα αποφάσεων μάθησης χρησιμοποιούν ένα εκπαιδευτικό σύνολο. Ο στόχος είναι να μάθουμε τη σχέση μεταξύ των προγνωστικών και της μεταβλητής στόχου στα δεδομένα εκπαίδευσης, έτσι ώστε το δέντρο να μπορεί να χρησιμοποιηθεί για την πρόβλεψη της ετικέτας κλάσης για νέα, αόρατα δεδομένα. (Elmachtoub, Liang and McNellis, 2020).

Η ακρίβεια ενός δέντρου αποφάσεων μπορεί να μετρηθεί από την ικανότητά του να προβλέπει σωστά τα αποτελέσματα των νέων δεδομένων. Αυτή η ακρίβεια τυπικά ποσοτικοποιείται χρησιμοποιώντας κάποιο είδος μέτρησης σφάλματος, όπως το μέσο τετράγωνο σφάλματος ή το ποσοστό σφάλματος ταξινόμησης.

Τα δέντρα αποφάσεων μπορούν να δημιουργηθούν χρησιμοποιώντας διάφορους αλγόριθμους, όπως ο αλγόριθμος ID3 ή ο αλγόριθμος C4.5. Αυτοί οι αλγόριθμοι αποφασίζουν πώς θα χωριστεί το δέντρο σε κάθε κόμβο. Ακόμα χρησιμοποιούν το κέρδος πληροφοριών, το οποίο είναι ένα μέτρο του πόσες «πληροφορίες» αποκτώνται λαμβάνοντας υπόψη ένα συγκεκριμένο χαρακτηριστικό (Elmachtoub, Liang and McNellis, 2020). Μερικά πλεονεκτήματα

των δέντρων αποφάσεων περιλαμβάνουν, διαφάνεια, προβλεψιμότητα, αντοχή στην υπερπροσαρμογή. Ακόμη και για κάποιον χωρίς εμπειρία στην επιστήμη δεδομένων ή στον προγραμματισμό, τα δέντρα αποφάσεων είναι εύκολα κατανοητά και οπτικά (Sosa Hernández *et al.*, 2021).

### **03. K – πλησιέστερου γείτονα (K-means)**

Ο K – πλησιέστερου γείτονα είναι ένας αλγόριθμος ταξινόμησης που βασίζεται στη μέτρηση της απόστασης μεταξύ χαρακτηριστικών διαφορετικών τιμών. Ο αλγόριθμος είναι εύκολο να γίνει αντιληπτός, είναι μη ευαίσθητος στα ακραία σημεία και κατάλληλος για ταξινομήσεις πολλών κλάσεων. Μειονέκτημα του είναι ότι είναι πολύ χρονοβόρος όταν χρησιμοποιείται για μεγάλα σύνολα δεδομένων (Tang *et al.*, 2020).

Χρησιμοποιείται ευρέως για την ταξινόμηση ενός συνόλου δεδομένων χωρίς ετικέτα σε διαφορετικές ομάδες. Το k αντιπροσωπεύει τον αριθμό των επιθυμητών συστάδων. Η αντικειμενική συνάρτηση του k-means αντιπροσωπεύει την απόσταση μεταξύ δεδομένων και συναφών κεντροειδών. Τα k-means τείνουν να εκχωρούν κάθε δεδομένο σε ένα σύμπλεγμα με το κεντροειδές που βρίσκεται πλησιέστερα στα δεδομένα. Η διαδικασία ενημέρωσης του κεντροειδούς με βάση ένα καθορισμένο σημείο δεδομένων θα επαναληφθεί έως ότου δεν αλλάξει κανένα σημείο ή το κεντροειδές. Η επιλογή του K έχει μεγάλο αντίκτυπο στην απόδοση του αλγορίθμου (Lovén *et al.*, 2019).

### **04. Τυχαία δάση (Random Forest)**

Το τυχαίο δάσος είναι ένα είδος αλγορίθμου δέντρων αποφάσεων, που δημιουργεί ένα σύνολο δέντρων αποφάσεων, στο οποίο κάθε δέντρο είναι υποδέντρο του μεγαλύτερου δάσους. Ο αλγόριθμος τυχαίου δάσους είναι ένας εποπτευόμενος αλγόριθμος μηχανικής μάθησης που είναι εξαιρετικά δημοφιλής και χρησιμοποιείται για προβλήματα ταξινόμησης και παλινδρόμησης στη Μηχανική μάθηση. Αποτελείται συνήθως από δέντρα πολλαπλών αποφάσεων. Για τον μετριασμό των υπερβολικών ζητημάτων του δέντρου αποφάσεων, η μέθοδος επιλέγει τυχαία ένα υποσύνολο χαρακτηριστικών για την κατασκευή κάθε δέντρου αποφάσεων. Ένα νέο σύνολο δεδομένων ταξινομείται από κάθε δέντρο αποφάσεων και στη συνέχεια το δείγμα δεδομένων κατηγοριοποιείται σε μια κατηγορία που συμφωνείται από τα περισσότερα δέντρα (Hoydis *et al.*, 2020).

Οι αλγόριθμοι τυχαίων δασών έχουν τρεις κύριες υπερπαραμέτρους, οι οποίες πρέπει να ρυθμιστούν πριν από τη χρήση: το μέγεθος του κόμβου, τον αριθμό των δέντρων και τον

αριθμό των χαρακτηριστικών που ελήφθησαν δειγματοληπτικά. Μετά ο τυχαίος ταξινομητής δασών μπορεί να χρησιμοποιηθεί για την επίλυση προβλημάτων παλινδρόμησης ή ταξινόμησης. Κάθε δέντρο στο σύνολο αποτελείται από ένα δείγμα δεδομένων που προέρχεται από ένα σετ εκπαίδευσης με αντικατάσταση, που ονομάζεται δείγμα εκκίνησης. Ανάλογα με τον τύπο του προβλήματος, ο προσδιορισμός της πρόβλεψης θα ποικίλλει. Για μια εργασία παλινδρόμησης, θα υπολογιστεί ο μέσος όρος των μεμονωμένων δέντρων απόφασης και για μια εργασία ταξινόμησης, η πλειοψηφία - δηλαδή η πιο συχνή κατηγορική μεταβλητή - θα δώσει την προβλεπόμενη κλάση (Schonlau and Zou, 2020).

## **05. Νευρωνικό δίκτυο (Neural Network)**

Το νευρωνικό δίκτυο μπορεί να χρησιμοποιηθεί για την εκμάθηση βιωματικών γνώσεων από ιστορικά δεδομένα από μεγάλο αριθμό μονάδων επεξεργασίας, οι οποίες λειτουργούν παράλληλα. Σε αυτές τις μονάδες συνήθως εφαρμόζονται οι συναρτήσεις ενεργοποίησης, όπως το σιγμοειδές, για την πραγματοποίηση μη γραμμικών υπολογισμών. Ένα νευρικό δίκτυο έχει συνήθως ένα επίπεδο εισόδου, ένα επίπεδο εξόδου και ένα ή περισσότερα κρυφά επίπεδα. Ρυθμίζοντας τον αριθμό των κρυφών επιπέδων και τον αριθμό των μονάδων σε κάθε επίπεδο, μπορούν να εκπαιδευτούν διαφορετικά μοντέλα για την επίλυση προβλημάτων ταξινόμησης ή παλινδρόμησης (Nawaz *et al.*, 2019). Ακόμα, τα νευρωνικά δίκτυα μπορούν να εφαρμοσθούν και για μη επιβλεπόμενη μάθηση (Αργυράκης, 2001).

Τα τεχνητά νευρωνικά δίκτυα αποτελούνται από στρώματα κόμβου, που περιέχουν ένα στρώμα εισόδου, ένα ή περισσότερα κρυφά στρώματα και ένα στρώμα εξόδου. Κάθε κόμβος, ή τεχνητός νευρώνας, συνδέεται με έναν άλλο. Εάν η έξοδος οποιουδήποτε μεμονωμένου κόμβου είναι πάνω από την καθορισμένη τιμή, αυτός ο κόμβος ενεργοποιείται στέλνοντας δεδομένα στο επόμενο επίπεδο του δικτύου, διαφορετικά δεν διαβιβάζονται δεδομένα στο επόμενο επίπεδο του δικτύου.

Καθώς τα δεδομένα μεταφέρονται από τη μια μονάδα στην άλλη, το νευρωνικό δίκτυο μαθαίνει όλο και περισσότερα για τα δεδομένα που τελικά καταλήγουν σε μια έξοδο από το επίπεδο εξόδου (Yang and Wang, 2020). Ακόμα περιέχουν τεχνητούς νευρώνες που ονομάζονται μονάδες. Αυτές οι μονάδες είναι διατεταγμένες σε μια σειρά επιπέδων που μαζί αποτελούν ολόκληρο το Τεχνητό Νευρωνικό Δίκτυο σε ένα σύστημα.

## **06. Μπαϊεζιανή Μάθηση**

Τα νευρωνικά δίκτυα Bayes αναφέρονται στην επέκταση τυπικών δικτύων με μεταγενέστερο συμπέρασμα προκειμένου να ελεγχθεί η προσαρμογή. Από μια ευρύτερη προοπτική, η Μπεϋζιανή προσέγγιση χρησιμοποιεί τη στατιστική μεθοδολογία, έτσι ώστε τα πάντα να έχουν μια κατανομή πιθανοτήτων συνδεδεμένη με αυτήν, συμπεριλαμβανομένων των παραμέτρων του μοντέλου στα νευρωνικά δίκτυα. Στις γλώσσες προγραμματισμού, οι μεταβλητές που μπορούν να λάβουν μια συγκεκριμένη τιμή θα μετατρέπουν το ίδιο αποτέλεσμα κάθε φορά που υπάρχει πρόσβαση στη συγκεκριμένη μεταβλητή (Wang and Yeung, 2021).

## **07. Μηχανές διανυσμάτων υποστήριξης**

Η πιο δημοφιλής τεχνική που σχετίζεται με τη Μηχανική μάθηση είναι η χρήση των Υποστήριξης Διανυσματικών Μηχανών (SVM). Αυτή είναι μια μέθοδος μάθησης που στοχεύει να βρει τον βαθμό διαχωρισμού μεταξύ δύο χαρακτηριστικών/κλάσεων (Pisner and Schnyer, 2020). Τα SVM χρησιμοποιούν δυαδικές τιμές για σύγκριση. Στην αρχή, αν για παράδειγμα τα δεδομένα είναι σε επίπεδο, μεταφέρονται σε μεγάλο μέγεθος και αφού αλλάξει η κλίμακα, επιλέγεται ένας πίνακας για να χωριστούν τα δεδομένα σε δύο ομάδες με τον καλύτερο τρόπο. Η ιδέα είναι να χαρτογραφηθούν τα αρχικά σχήματα σε έναν χώρο μεγαλύτερο από τον αρχικό, στην οποία τα γεγονότα μπορούν να διαχωριστούν (Cervantes *et al.*, 2020).

Ακόμα, διευκολύνουν την ταξινόμηση χαρακτηριστικών που δεν μπορούν να διαχωριστούν και αποτελούν σαφή αναπαράσταση δεδομένων. Χρησιμοποιούν έναν αλγόριθμο εκπαίδευσης ώστε να αντιπροσωπεύουν τις μη γραμμικές σχέσεις στο χώρο. Για τη μετατροπή από ένα διάστημα σε μεγαλύτερο χώρο χρησιμοποιούνται συναρτήσεις πυρήνα. Οι δύο πιο κοινές συναρτήσεις είναι το πολυώνυμο του πυρήνα και η συνάρτηση ακτινικής βάσης (rbf).

## **08. Αυτο-οργανωμένος χάρτης (SOM)**

Χρησιμοποιείται συχνά για την πραγματοποίηση μείωσης διαστάσεων και ομαδοποίησης δεδομένων. Περιέχει ένα επίπεδο εισόδου και ένα επίπεδο χάρτη. Κάθε επίπεδο περιλαμβάνει πολλούς νευρώνες και κάθε νευρώνας έχει φορέα βάρους. Κατά τη διάρκεια της εκπαιδευτικής διαδικασίας, μπορεί να δημιουργήσει και να αναδιοργανώσει τον χάρτη.

Σε αντίθεση με τα συμβατικά νευρωνικά δίκτυα που εφαρμόζουν μάθηση διόρθωσης σφαλμάτων, οι αυτό-οργανωμένοι χάρτες χρησιμοποιούν μια μη εποπτευόμενη προσέγγιση ανταγωνιστικής μάθησης. Μετά την εκπαίδευση, ένας νέος φορέας εισόδου ταξινομείται σε ένα σύμπλεγμα με βάση τον νευρώνα που κερδίζει στον χάρτη. Οι τεχνικές έχουν

χρησιμοποιηθεί με επιτυχία σε διάφορες εργασίες αναγνώρισης προτύπων (Zhang et al., 2019).

## 09. Κρυφό μοντέλο Markov (HMM)

Η προσέγγιση μπορεί να χρησιμοποιηθεί για τη μοντελοποίηση ενός συστήματος με μια διαδικασία Markov με άγνωστες παραμέτρους. Η κύρια πρόκληση είναι ο προσδιορισμός των κρυφών παραμέτρων από γνωστές παραμέτρους. Τα μοντέλα Markov χρησιμοποιούνται ευρέως σε τυχαία δυναμικά σενάρια με ιδιότητα χωρίς μνήμη, πράγμα που σημαίνει ότι η υπό όρους πιθανότητα κατανομής μελλοντικών καταστάσεων εξαρτάται μόνο από την τρέχουσα κατάσταση. Οι παράμετροι του μπορούν να εκπαιδευτούν με εποπτεία ή χωρίς. Μπορεί να χρησιμοποιηθεί για την αναπαράσταση μη στατικών αλληλουχιών, η οποία επιτρέπει στο σύστημα να αλλάζει με την πάροδο του χρόνου με διαφορετικές κατανομές πιθανότητας εξόδου από κάθε κατάσταση (Du et al., 2020).

### 1.5 Πλεονεκτήματα και μειονεκτήματα της μηχανικής μάθησης

Η μηχανική μάθηση αντιπροσωπεύει έναν σημαντικό κλάδο της τεχνητής νοημοσύνης που έχει επαναπροσδιορίσει τον τρόπο με τον οποίο αντιλαμβανόμαστε και αναλύουμε τα δεδομένα. Μέσω αλγορίθμων και μοντέλων που εκπαιδεύονται σε μεγάλα σύνολα δεδομένων, η μηχανική μάθηση επιτρέπει στις μηχανές να "μαθαίνουν" από τα δεδομένα και να προβλέπουν, να αναγνωρίζουν πρότυπα και να λαμβάνουν αποφάσεις χωρίς να χρειάζεται να προγραμματιστούν ειδικά για αυτό. Αυτή η τεχνολογία έχει επαναπροσδιορίσει τον τρόπο λειτουργίας πολλών βιομηχανικών κλάδων και διαδικασιών, παρέχοντας πληθώρα πλεονεκτημάτων. Στον κλάδο της λογιστικής και του ελέγχου, η εφαρμογή της μηχανικής μάθησης, σύμφωνα με σύντομη παρουσίαση των βασικότερων γεγονότων της προηγούμενης βιβλιογραφίας, παρουσιάζει τα εξής πλεονεκτήματα (Dineva and Atanasova, 2020):

- Αυξημένη αποδοτικότητα και αυτοματοποίηση: Η τεχνολογία μηχανικής μάθησης μπορεί να επεξεργαστεί γρήγορα και αποτελεσματικά μεγάλες ποσότητες δεδομένων, επιτρέποντας ταχύτερες και ακριβέστερες αποφάσεις. Αυτό μπορεί να οδηγήσει σε αυξημένη παραγωγικότητα και εξοικονόμηση κόστους για τις εταιρείες.
- Βελτιωμένες δυνατότητες λήψης αποφάσεων: Οι αλγόριθμοι μηχανικής μάθησης μπορούν να εντοπίζουν μοτίβα και να κάνουν προβλέψεις με βάση ιστορικά δεδομένα, γεγονός που μπορεί να βελτιώσει την ακρίβεια των αποφάσεων και να μειώσει την ανάγκη για ανθρώπινη επίβλεψη.

- Δυνατότητα χειρισμού μεγάλου όγκου δεδομένων: Οι αλγόριθμοι μηχανικής μάθησης μπορούν να χειριστούν και να αναλύσουν μεγάλες ποσότητες δεδομένων, γεγονός που μπορεί να είναι χρήσιμο για οργανισμούς με τεράστιες ποσότητες πληροφοριών προς επεξεργασία.
- Δυνατότητα εξοικονόμησης κόστους: Με την αυτοματοποίηση των διαδικασιών και τη βελτίωση της λήψης αποφάσεων, η Μηχανική Μάθηση μπορεί να εξοικονομήσει χρήματα στις εταιρείες.
- Δυνατότητα εντοπισμού μοτίβων και πραγματοποίησης προβλέψεων: Μέσω της Μηχανικής Μάθησης, μπορεί να εντοπιστούν μοτίβα και να γίνουν προβλέψεις με βάση ιστορικά δεδομένα, γεγονός που μπορεί να είναι χρήσιμο για προβλέψεις, ανίχνευση ανωμαλιών και άλλες εργασίες που απαιτούν τον εντοπισμό μοτίβων στα δεδομένα. Παρά τα αξιοσημείωτα πλεονεκτήματα που προσφέρει η μηχανική μάθηση, υπάρχουν επίσης και ορισμένα σημαντικά μειονεκτήματα που πρέπει να ληφθούν υπόψη (Dineva and Atanasova, 2020):
- Υψηλό κόστος εφαρμογής: Απαιτείται υλικό, λογισμικό και εξειδικευμένο προσωπικό με τις απαραίτητες δεξιότητες για το σχεδιασμό και τη συντήρηση του συστήματος.
- Ανάγκη για μεγάλο όγκο δεδομένων: Η επεξεργασία της μηχανικής μάθησης απαιτεί μεγάλα δεδομένα για να είναι αποτελεσματική, γεγονός που μπορεί να αποτελέσει πρόκληση για οργανισμούς με περιορισμένα σύνολα δεδομένων.
- Εξάρτηση από την ποιότητα των δεδομένων: Η αποτελεσματικότητα της Μηχανικής Μάθησης και της τεχνητής νοημοσύνης εξαρτάται σε μεγάλο βαθμό από την ποιότητα των δεδομένων που χρησιμοποιούνται για την εκπαίδευσή της. Η κακή ποιότητα δεδομένων μπορεί να οδηγήσει σε ανακριβή ή αναξιόπιστα αποτελέσματα.
- Μεροληψία στα δεδομένα ή τους αλγορίθμους: Οι αλγόριθμοι μηχανικής μάθησης μπορεί να εισάγουν ακούσια μεροληψία εάν τα δεδομένα που χρησιμοποιούνται για την εκπαίδευσή τους δεν είναι αντιπροσωπευτικά του πληθυσμού-στόχου. Επιπλέον, οι ίδιοι οι αλγόριθμοι μπορεί να διακρίνουν τυχόν προκαταλήψεις που υπάρχουν στα δεδομένα.
- Έλλειψη διαφάνειας στη λήψη αποφάσεων: Τα προγράμματα μηχανικής μάθησης μπορεί να είναι δύσκολο να κατανοηθούν και να ερμηνευθούν, καθιστώντας δύσκολο να εξηγηθεί πώς κατέληξαν σε μια συγκεκριμένη απόφαση. Αυτή η έλλειψη διαφάνειας

μπορεί να προβληματίσει τους οργανισμούς που πρέπει να λογοδοτήσουν για τις διαδικασίες λήψης αποφάσεων.

- Πιθανή απώλεια θέσεων εργασίας: Καθώς τα συστήματα μηχανικής μάθησης εκτελούν εργασίες που προηγουμένως εκτελούνταν από ανθρώπους, υπάρχει κίνδυνος απώλειας θέσεων εργασίας για όσους εκτελούσαν προηγουμένως αυτές τις εργασίες.

#### 1.6 Μηχανική Μάθηση ως Αντικείμενο Μελέτης

Σε γενικές γραμμές οι μηχανές μπορούν να αποκτήσουν ευφυΐα που τους επιτρέπει να μάθουν από προηγούμενους υπολογισμούς και να προσαρμοστούν στο περιβάλλον, επειδή έχει αναπτυχθεί η μηχανική μάθηση. Ο σκοπός της μηχανικής μάθησης είναι να ανιχνεύει μοτίβα και να μαθαίνει πώς να κάνει προβλέψεις και συστάσεις μέσω της επεξεργασίας δεδομένων.

Ως αντικείμενο μελέτης, η μηχανική μάθηση περιλαμβάνει την κατανόηση των υποκείμενων αρχών και εννοιών του πεδίου, συμπεριλαμβανομένων διαφόρων αλγορίθμων και τεχνικών για την εκπαίδευση μοντέλων, την αξιολόγηση της απόδοσής τους και την ανάπτυξή τους σε πραγματικές εφαρμογές. Περιλαμβάνει ακόμα την ανάπτυξη ισχυρών θεμελίων στα μαθηματικά, τη στατιστική και την επιστήμη των υπολογιστών, τα οποία είναι απαραίτητα για την κατανόηση και την ανάπτυξη αλγορίθμων μηχανικής μάθησης (Πειβάνη, 2020). Συνολικά, η μελέτη της μηχανικής μάθησης μπορεί να αποτελέσει μια ικανοποιητική και προκλητική εμπειρία που μπορεί να ανοίξει συναρπαστικές ευκαιρίες σταδιοδρομίας σε ένα ευρύ φάσμα κλάδων, όπως η υγειονομική περίθαλψη, η χρηματοδότηση, το ηλεκτρονικό εμπόριο και άλλα. Η μηχανική μάθηση έχει εξελιχθεί σε ερευνητικό θέμα, καθώς επιτρέπει στους χρήστες να αποκτήσουν βαθύτερες γνώσεις και προβλέψεις σχετικά με την εξέλιξη του φαινομένου που μελετάται και την αποτελεσματική χρήση των υπάρχοντων δεδομένων (Γεωργούλη, 2015).



## Κεφάλαιο 2<sup>ο</sup>: Φορολογία και Φοροδιαφυγή

Η φορολογία αποτελεί ένα θεμέλιο στοιχείο της οικονομικής λειτουργίας κάθε κράτους, παρέχοντας τους απαραίτητους πόρους για τη χρηματοδότηση των δημοσίων υπηρεσιών και των υποδομών. Ωστόσο, η φοροδιαφυγή αντιπροσωπεύει ένα σημαντικό πρόβλημα για τις φορολογικές αρχές και την οικονομία γενικότερα. Η φοροδιαφυγή αναφέρεται στις πρακτικές όπου οι φορολογούμενοι αποφεύγουν την πλήρη ή τμηματική καταβολή των φορολογικών τους υποχρεώσεων, είτε με νόμιμους είτε με παράνομους τρόπους.

Η φοροδιαφυγή έχει διάφορες μορφές, συμπεριλαμβανομένης της απόκρυψης εσόδων, της υπερεκτίμησης εξόδων, της μεταφοράς κεφαλαίων σε φορολογικά φιλικές δικαιοδοσίες, και της χρήσης νομικών τεχνικών για να μειωθεί η φορολογική επιβάρυνση. Μπορεί να έχει σοβαρές συνέπειες, όπως τη μείωση των δημοσίων εσόδων, την αύξηση του δημοσίου χρέους και την αναστάτωση της δικαιότερης κατανομής των επιβαρύνσεων στην κοινωνία.

### 2.1 Ιστορική εξέλιξη φορολογίας

Η έννοια της φορολογίας υπήρχε από τα αρχαία χρόνια. Η φορολογία φαίνεται να αρχίζει από την πρώτη δυναστεία του Παλαιού Βασιλείου της Αρχαίας Αιγύπτου, 3000-2800 π.Χ. Κατά την περίοδο αυτή υπάρχουν δύο είδη φόρων η καταναγκαστική εργασία (corvée) και η δεκάτη. Η καταναγκαστική εργασία (corvée) ήταν μια κατάσταση κατά την οποία επιβαλλόταν η εργασία στους κατοίκους της υπαίθρου που ήταν φτωχοί και δεν μπορούσαν να πληρώσουν τους φόρους. Ακόμα, σε ένα corvée δεν ήταν αναγκαίο ο πληθυσμός να έχει γη, καλλιέργειες ή μετρητά (Brown, 1994). Η δεκάτη ήταν ένας τακτικός φόρος για την αγροτική παραγωγή, ίσος με το 10% της αξίας των προϊόντων της συνολικής ακαθάριστης παραγωγής, και εισπραττόταν σε είδος ή χρήμα. Ο φόρος της δεκάτης υπήρχε μέχρι το 1880. Μια μετέπειτα μορφή του φόρου αυτού είναι ο φόρος εισοδήματος.

Ωστόσο, σύμφωνα με τον Adams, οι πρώτες αναφορές ενός οργανωμένου φορολογικού συστήματος, με σαφές οργανωτικό πλαίσιο κατά της φοροδιαφυγής, χρονολογούνται από τον αιγυπτιακό πολιτισμό, όπου οι Φαραώ καθιέρωσαν έναν μηχανισμό με επίκεντρο τους ακριβοπληρωμένους φοροεισπράκτορες, γνωστούς ως γραφείς (scribes) (Adams, 2006). Αξίζει να σημειωθεί ότι επί τον Φαραώ επιβάλλονταν φόροι στο μαγειρικό λάδι, κατά το οποίο οι γραφείς έλεγχαν τα νοικοκυριά ώστε να βεβαιωθούν για την ποσότητα του μαγειρικού ελαίου

που κατανάλωναν και ότι δεν χρησιμοποιούσαν υπολείμματα που δημιουργούνταν από άλλες διαδικασίες μαγειρέματος ως υποκατάστατα για την επιβολή φόρου στο μαγειρικό λάδι.

Λίγο αργότερα κατά την ομηρική περίοδο υπήρχαν δύο είδη άμεσων φόρων, οι «δοτιναί» που ονομάζονταν τα δώρα που δίνονταν υπό ειδικούς όρους και οι «θέμιστες» που ήταν δωρεές που έπρεπε να καταβάλλονται περιοδικά.

Επιπρόσθετα, ο Adams υποστηρίζει ότι οι Αθηναίοι έβαζαν πολύ υψηλούς φόρους στην περιουσία και τον πλούτο σε περιόδους πολέμου, που ονομάζονταν εισφορές, και κανείς δεν μπορούσε να εξαιρεθεί από την καταβολή τους, και όσοι δεν πλήρωναν τους φόρους θα τιμωρούνταν ακόμα μέχρι και σε θάνατο. Επιπλέον, τονίζει ότι ο ελληνικός πολιτισμός ήταν ένας από τους λίγους πολιτισμούς που πέτυχαν να ανατρέψουν την υπάρχουσα φορολογική πολιτική όταν το απαιτούσε η έκτακτη ανάγκη και να καταργήσουν την έκτακτη φορολογία όταν εξαλείφονταν η έκτακτη ανάγκη. Στην πραγματικότητα, οι περισσότεροι από τους σημερινούς σύγχρονους φόρους δικαιολογούνται σε περιόδους πολέμου, ωστόσο, η αδυναμία και η απροθυμία των κυβερνήσεων να τους καταργήσουν μετά το τέλος του πολέμου οδήγησε στη μόνιμη είσπραξή τους (Adams, 2006).

Στη Ρωμαϊκή Αυτοκρατορία, υπήρξε για πρώτη φορά φόρος εισαγωγής και εξαγωγής, που ονομαζόταν «πορτόρια». Ο Αύγουστος και ο Ιούλιος Καίσαρας ήταν φορολογικοί σχεδιαστές σε όλη τη διάρκεια της βασιλείας τους. Ο πρώτος εισήγαγε φόρο κληρονομιάς, ενώ ο δεύτερος ήταν υπεύθυνος για τον φόρο επί των πωλήσεων 1%, που σήμερα θεωρείται το ισοδύναμο του φόρου προστιθέμενης αξίας (ΦΠΑ). Αυτός ο φόρος περιλαμβάνει την υποχρέωση φορολογίας όλων των μορφών πλούτου, συμπεριλαμβανομένων της γης, των σπιτιών, των σκλάβων, των ζώων, των χρημάτων και των προσωπικών αντικειμένων. Στο σημείο αυτό είναι αναγκαίο να σημειωθεί ότι, αν και αυτό το ποσοστό είναι μικρό η ύπαρξη φοροδιαφυγής επιβεβαιώνει τη διαιώνιση ενός σύγχρονου προβλήματος, το οποίο έπρεπε να αντιμετωπίσει και η Ρωμαϊκή Αυτοκρατορία.

Όσο αφορά τον ελλαδικό χώρο, επί Τουρκοκρατίας, ο κύριος φόρος ήταν το λεγόμενο «χαράτσι», που πλήρωναν όλοι οι άντρες ηλικίας 15 ετών και άνω. Στις περιπτώσεις που κάποιος υπόδουλος πατέρας δήλωνε μικρότερη ηλικία του τέκνου του προκειμένου να αποφύγει την καταβολή φόρου ξεγελώντας τους εντεταλμένους είσπραξης των φόρων, αυτοί χρησιμοποιώντας ένα σχοινί, του οποίου το μήκος καθόριζαν κατά το δοκούν, ως μέτρο μέτρησης της οργιάς, υπολόγιζαν από το μέγεθος του κεφαλιού του την ηλικία του παιδιού και επέβαλαν την φορολογία. Η δεκάτη είναι ένας φόρος ίσος με το 10% της αξίας του

συνολικού παραγόμενου προϊόντος, που επιβάλλεται σε είδος ή σε νόμισμα. Ήταν ένας φόρος αγροτικής παραγωγής που επιβαλλόταν στους χριστιανούς σκλάβους από τους Οθωμανούς κατακτητές. Η δεκάτη διατηρήθηκαν μέχρι το 1880.

Στις Ηνωμένες Πολιτείες της Αμερικής, μετά τον πόλεμο κατά της Μεγάλης Βρετανίας (1775-1783), η νέα κυβέρνηση ήταν απρόθυμη να φορολογήσει ορισμένα αγαθά, όπως το αλκοόλ, τον καπνό και τη ζάχαρη, κυρίως λόγω φορολογικών ζητημάτων. Αργότερα επιβλήθηκαν φόροι λόγω της ανάγκης χρηματοδότησης των πολέμων. Η Γαλλία εισήγαγε έναν φόρο κληρονομιάς το 1790, και εξακολουθεί να υπάρχει σήμερα. Μετά τον Αμερικανικό Εμφύλιο Πόλεμο το 1872 και την καταστροφή που προκάλεσε, το Κογκρέσο ψήφισε τον νόμο περί εσόδων του 1861, εισάγοντας τη Διεθνή Υπηρεσία Εσόδων (IRS), η οποία αποτελεί τον μηχανισμό που καθιέρωσε το σύγχρονο αμερικανικό φορολογικό σύστημα.

Η επιβολή φορολογίας στην νεότερη Ελλάδα έκανε την εμφάνισή της με το ψήφισμα της 4ης Φεβρουαρίου του 1830 ως φόρος επιτηδεύματος, ενώ εδραιώθηκε με τον νόμο ΧΚ' 620 του 1877, όπου επιβλήθηκε φόρος εισοδήματος στο μέρισμα που διένειμαν στους μετόχους τους οι ανώνυμες εταιρίες. Η φορολογία των κινητών αξιών ενσωματώθηκε στην ελληνική φορολογική νομοθεσία ως διάταξη του νόμου «Περί Χαρτοσήμου» και θεσμοθετήθηκε σε έκτακτη κοινοβουλευτική περίοδο, το 1877 από την τότε οικουμενική κυβέρνηση του Κανάρη. Αντικείμενο φορολόγησης είχε το διανεμόμενο μέρισμα των ανωνύμων εταιριών και το πραγματοποιούμενο καθαρό κέρδος των ξένων ανωνύμων εταιριών, που είχαν αναπτύξει επιχειρηματική δραστηριότητα στην Ελλάδα. Μέχρι την περίοδο αυτή οι ανώνυμες εταιρείες και κατ' επέκταση οι χρηματοπιστωτικές επιχειρήσεις, όπως η Εθνική Τράπεζα Ελλάδος πλήρωναν φόρο επιτηδεύματος.

Τον 20ο αιώνα, δύο νομοθετικά διατάγματα ήρθαν στο προσκήνιο της φορολογίας και έγιναν ορόσημα της φορολογίας εισοδήματος. Συγκεκριμένα, το Νομοθετικό Διάταγμα (Ν.Δ) 3323/1955 ρυθμίζει τη φορολογία των φυσικών προσώπων, θεσπίζοντας προοδευτικό φόρο στο ακαθάριστο εισόδημα των φυσικών προσώπων, ενώ το Ν.Δ. 3843/1958 ρυθμίζει τη φορολογία των νομικών προσώπων. Το διάταγμα αναφέρει ότι «οι ημεδαπές εταιρείες φορολογούνται στο όνομά τους και επί των αδιανέμητων κερδών, ενώ τα κέρδη που διανέμονται ως μερίσματα στους μετόχους ή ως αμοιβές και ποσοστά μελών διοικητικών συμβουλίων ή άλλες αμοιβές.

Ακολούθησε ο Νόμος 2065/1992, σύμφωνα με τον οποίο φορολογούνται όλα τα φορολογητέα κέρδη των ανωνύμων εταιριών και ισχύει και για προσωπικές εταιρείες και εταιρείες

περιορισμένης ευθύνης, με ορισμένες αλλαγές. Τέλος, τα Νομοθετικά Διατάγματα (Ν.Δ. 3323/1955 και Ν.Δ. 3843/1958) έχουν πλέον κωδικοποιηθεί σε ενιαίο κείμενο, δυνάμει του Ν. 2238/1994 «Κυρώσεις του Νόμου Φορολογίας Εισοδήματος». Τέλος, βάσει αυτού του νόμου, ο νόμος περί φόρου εισοδήματος θεσπίστηκε το 1994 και έχει τροποποιηθεί αρκετές φορές σύμφωνα με τις ανάγκες του φορολογικού συστήματος της χώρας. Σήμερα τέθηκε σε ισχύ ο νέος Κώδικας Φορολογίας Εισοδήματος που θεσπίστηκε με το Ν. 4172/2013, ο οποίος, σε συνδυασμό με τον Ν. 4174/2013, Κώδικα Φορολογικών Διαδικασιών, αποτελεί μια μεταρρυθμιστική προσπάθεια με στόχο τον εκσυγχρονισμό της φορολογικής διοίκησης και την εναρμόνισή της με τις νέες τεχνολογίες και τους νέους στόχους της δημοσιονομικής πολιτικής.

## 2.2 Φορολογία

Οι δημόσιοι φορείς χρειάζονται πόρους για να παρέχουν δωρεάν υπηρεσίες και αγαθά στο κοινωνικό σύνολο, καλύπτοντας παράλληλα το κόστος τους. Η κύρια πηγή χρηματοδότησης του κράτους για την κάλυψη των αναγκών είναι η φορολογία. Η φορολογία λοιπόν, είναι ένα μέσο για την εξαναγκαστική μεταφορά πόρων από τον ιδιωτικό στον δημόσιο τομέα, ώστε να μπορεί να χρηματοδοτήσει το κόστος των υπηρεσιών που παρέχει στο κοινωνικό σύνολο δωρεάν (Γεωργακόπουλος, 2016). Στη συνέχεια ακολουθεί αναφορά στις βασικές έννοιες που διέπουν ένα φορολογικό σύστημα, στις βασικές λειτουργίες των φόρων, στους κανόνες ενός φορολογικού συστήματος και ποιος τελικά είναι ο σκοπός της φορολογίας.

### 2.2.1 Βασικές έννοιες

Σύμφωνα με το λεξικό του Τεγόπουλου – Φυτράκη, η λέξη «φορολογία» είναι μια σύνθετη λέξη <φόρος + λόγος> και σημαίνει η επιβολή φόρου, η υποχρέωση φυσικών ή νομικών προσώπων να καταβάλλουν φόρο στο δημόσιο (Τεγόπουλος and Φυτράκης, 1990). Πιο αναλυτικά, ο «φόρος» είναι μία έννοια στην οποία έχουν δοθεί πολλοί ορισμοί, καθώς εμφανίζεται ως λέξη από την αρχαιότητα. Προέρχεται από τη λατινική λέξη *taxo*, που σημαίνει “rate”, είναι ένα τέλος που επιβάλλεται από μια κυβέρνηση πάνω σε ένα προϊόν, εισόδημα ή δραστηριότητα. Σύμφωνα με το Λεξικό της κοινής νεοελληνικής (1998) ο «φόρος» [λόγ. < αρχ. φόρος] χωρίζεται σε άμεσος, που καταβάλλεται απευθείας από τον οφειλέτη στο κράτος και έμμεσος, που περιέχεται στην τιμή προϊόντων ή υπηρεσιών (Ίδρυμα Τριανταφυλλίδη, 1998). Επομένως, υπάρχουν διάφορες μορφές φόρου. Μπορεί να προέρχεται είτε από τη φορολογία εισοδήματος και περιουσιακών στοιχείων, είτε από αγορά προϊόντων και υπηρεσιών (ΦΠΑ). Ο φόρος που καταβάλλει ο κάθε πολίτης είναι διαφορετικός και υπολογίζεται αναλογικά με τις ενέργειες του καθώς και με την περιουσία και τους ίδιους

πόρους που διαθέτει. Φόρο αποδίδουν οι υπήκοοι του εκάστοτε κράτους, αλλά και πρόσωπα που ζουν ή εργάζονται σε αυτό, δηλαδή αλλοδαποί, μόνιμα ή προσωρινά εγκατεστημένοι.

Η φορολογία στη σύγχρονη κοινωνία χαρακτηρίζεται από δύο διαστάσεις, την οικονομική και νομική. Ο γενικός ορισμός της έννοιας του φόρου που δίνεται από τη Γερμανική Φορολογική Νομοθεσία είναι: «Οι φόροι είναι χρηματικές παροχές που δεν αποτελούν μια ανταποδοτική ενέργεια σε κάποια συγκεκριμένη παροχή και επιβάλλονται από ένα όργανο δημοσίου δικαίου σε όλους όσους συμπίπτει να έχουν την ίδια πραγματική κατάσταση, στην οποία ο νόμος αποδίδει την υποχρέωση παροχής. Η επιδίωξη εισόδων πιθανόν να τίθεται ως δεύτερος στόχος.»

Ο πλέον πιο διαδεδομένος ορισμός για τη φορολογία στα πλαίσια της νομικής επιστήμης αποδίδεται στον Gaston Jèze ο οποίος ισχυρίστηκε ότι «ο φόρος αποτελεί αναγκαστική εισφορά δίχως ειδική αντιπαροχή από τον πολίτη προς το κράτος για την αντιμετώπιση των δημοσίων δαπανών».

Τελειώνοντας, μια σωστή φορολογική πολιτική μπορεί να χρησιμοποιηθεί για την προώθηση διαφόρων οικονομικών στόχων. Οι φόροι επιμερίζονται στις διάφορες λειτουργικές δαπάνες του δημοσίου αλλά και στις δαπάνες των δραστηριοτήτων που παρέχουν οι δημόσιοι φορείς καθώς και στα δημόσια αγαθά που προσφέρουν στους πολίτες. Έτσι, μέσω των φόρων η κυβέρνηση χρηματοδοτεί τις δημόσιες δαπάνες της και επιτυγχάνει τους στόχους της σταθεροποίησης της οικονομίας και της δίκαιης διανομής και ανακατανομής του εισοδήματος.

### 2.2.2 Βασικές λειτουργίες των φόρων

Οι φόροι έχουν τρεις κύριες λειτουργίες (Γεωργακόπουλος, 2012). Η πρώτη από αυτές είναι η ταμειυτική λειτουργία μπορεί να είναι η πιο βασική λειτουργία της φορολογίας, αλλά δεν είναι η μόνη. Επιτυγχάνει τη συγκέντρωση των απαιτούμενων για την αντιμετώπιση των δημοσίων δαπανών.

Έτσι, ένας από τους βασικούς σκοπούς της δημοσιονομικής πολιτικής πρέπει να είναι η επάρκεια των φορολογικών εσόδων για την κάλυψη των δημοσίων δαπανών.

Εκτός από αυτό, οι φόροι έχουν και άλλες λειτουργίες καθώς χρησιμοποιούνται από τους δημόσιους φορείς ως εργαλεία για την εφαρμογή οικονομικών και κοινωνικών πολιτικών. Έτσι, συχνά γίνεται διάκριση μεταξύ της οικονομικής λειτουργίας της φορολογίας και της κοινωνικής λειτουργίας της φορολογίας.

Η κοινωνικοοικονομική λειτουργία αναφέρεται στη χρήση τους ως εργαλεία οικονομικής πολιτικής για τη σταθεροποίηση της οικονομίας σε πλήρη απασχόληση, για την καταπολέμηση της ανεργίας και του πληθωρισμού και για την προώθηση της οικονομικής ανάπτυξης μιας χώρας. Ο φόρος χρησιμοποιείται ως μέσο για την επίτευξη βασικών σκοπών από την κυβερνητική πολιτική οι οποίοι μπορεί να είναι πρώτον η οικονομική ανάπτυξη της χώρας, δεύτερον η επίτευξη της απασχόλησης και τρίτο η ανακατανομή των παραγωγικών πόρων με σκοπό την ανάπτυξη ορισμένων παραγωγικών κλάδων ή ορισμένων περιοχών της χώρας (περιφερειακή ανάπτυξη).

Τέλος υπάρχει η κοινωνικοπολιτική λειτουργία, αναφέρεται στη χρήση των φόρων ως μέσων άσκησης κοινωνικής πολιτικής, δηλαδή για την αναδιανομή του εισοδήματος και του πλούτου έτσι ώστε να ανταποκρίνεται στις αντιλήψεις της κοινωνίας περί κοινωνικής δικαιοσύνης.

Η εφαρμογή της αρχής αυτής σε συνδυασμό με τη δημιουργία δαπανών κοινωνικής πρόνοιας και τη λήψη μέτρων οικονομικής πολιτικής για την εξασφάλιση υψηλού επιπέδου απασχόλησης του παραγωγικού δυναμικού κάθε χώρας απετέλεσαν τα κυριότερα μέσα με τα οποία συντελέστηκε βαθμηδόν η άμβλυση ή εξάλειψη των βασικών μειονεκτημάτων του κεφαλαιοκρατικού οικονομικού συστήματος.

### 2.2.3 Κανόνες ενός φορολογικού συστήματος

Ο Adam Smith ήταν ο πρώτος που έκανε οριοθέτηση φορολογικών αρχών. Στο έργο του «The Wealth of Nations» (1776) διατυπώνει τους κανόνες της φορολογίας (cannons of taxation), που είναι απαραίτητο να έχει ένα φορολογικό σύστημα. Κάνει διάκριση μεταξύ δύο σχολών της δημόσιας οικονομικής: της Αγγλίας και Γαλλίας που βασίζεται στην πολιτική οικονομία και της Γερμανίας και της Αυστρίας που θεωρεί τη δημοσιολογία ως κλάδο του διοικητικού δικαίου, που ιστορικά την αποδίδει στις καμεραλιστικές σπουδές στις χώρες αυτές (Θεοχαράκης, 2018). Παρατηρεί μάλιστα ότι «Η εμφάνιση της πολιτικής οικονομίας εν Γερμανία εχαλάρωσε, προς στιγμήν, τας στενωπάτας ταύτας σχέσεις. Αλλ' η επιρροή του Αντάμ Σμιθ υπήρξεν ενταύθα παροδική» (Ανδρεάδης, 1903).

Οι τέσσερις κανόνες για τη φορολογία του Adam Smith είναι:

- 1)** Κανόνας της φορολογικής δικαιοσύνης και ισότητας: Οι φορολογούμενοι πρέπει να εισφέρουν στις δαπάνες του κράτους ανάλογα με τις δυνατότητές τους.
- 2)** Κανόνας της βεβαιότητας του φόρου: Ο φόρος να είναι βέβαιος και όχι αυθαίρετος. Ο χρόνος και ο τρόπος της πληρωμής του φόρου, καθώς και το πληρωτέο ποσό του, πρέπει να είναι σαφώς καθορισμένα.

- 3) Κανόνας της καταλληλότητας του φόρου: Οι φόροι πρέπει να καταβάλλονται σε χρόνο και με τρόπο που να είναι οι πιο κατάλληλοι για το φορολογούμενο.
- 4) Κανόνας οικονομικότητας του φόρου: Ο φόρος πρέπει να επιβάλλεται με τέτοιο τρόπο ώστε η διαφορά ανάμεσα στο φόρο που καταβάλουν οι πολίτες και το ποσό που εισπράττει το δημόσιο να είναι όσο το δυνατό μικρότερη.

#### 2.2.4 Σκοπός της φορολογίας

Κατά τον 19ο αιώνα η επικρατούσα ιδέα ήταν ότι οι φόροι έπρεπε να χρησιμεύουν κυρίως για τη χρηματοδότηση της κυβέρνησης. Σε παλαιότερες εποχές, και ξανά σήμερα, οι κυβερνήσεις χρησιμοποίησαν τη φορολογία για άλλους εκτός από απλώς φορολογικούς σκοπούς. Ένας χρήσιμος τρόπος για να δείτε τον σκοπό της φορολογίας, που αποδίδεται στον Αμερικανό οικονομολόγο. Ο Richard A. Musgrave, είναι να κάνει διάκριση μεταξύ των στόχων της κατανομής των πόρων, της αναδιανομής του εισοδήματος και της οικονομικής σταθερότητας. Η οικονομική ανάπτυξη ή ανάπτυξη και η διεθνής ανταγωνιστικότητα αναφέρονται μερικές φορές ως ξεχωριστοί στόχοι, αλλά γενικά μπορούν να υπαχθούν στους άλλους τρεις. Ελλείψει ισχυρού λόγου παρέμβασης, όπως η ανάγκη μείωσης της ρύπανσης, ο πρώτος στόχος, η κατανομή των πόρων, προωθείται εάν η φορολογική πολιτική δεν παρεμβαίνει σε κατανομές που καθορίζονται από την αγορά. Ο δεύτερος στόχος, η αναδιανομή του εισοδήματος, αποσκοπεί στη μείωση των ανισοτήτων στην κατανομή του εισοδήματος και του πλούτου. Ο στόχος της σταθεροποίησης, που εφαρμόζεται μέσω της φορολογικής πολιτικής, της πολιτικής για τις κρατικές δαπάνες, της νομισματικής πολιτικής και διαχείρισης του χρέους, αναφέρεται στη διατήρηση υψηλής απασχόλησης και σταθερότητας των τιμών .

Ταυτόχρονα, άλλες συνταγματικές διατάξεις όπως ο σεβασμός της ιδιοκτησίας, το καθήκον του κράτους να διασφαλίζει τη γενική ευημερία των πολιτών του, η προστασία και ο σεβασμός των ανθρωπίνων αξιών, το δικαίωμα στην ελεύθερη ανάπτυξη της προσωπικότητας και η αρχή της αναλογικότητας είναι επίσης απαραίτητες παράμετροι που πρέπει να λαμβάνονται υπόψη στην επιβολή φόρου, αλλά και στις ποινές για φοροδιαφυγή.

Στην περίπλοκη πλεξίδα της κοινωνίας, η φορολογία υφαίνει το αόρατο ύφασμα της δικαιοσύνης και της ισορροπίας. Είναι ο φόρος που επιβάλλεται στην απόκτηση, στην εργασία, στην περιουσία, ενώ ταυτόχρονα αντιτίθεται ως ένας φόρος που πληρώνεται για τις αξίες μας, για τον τρόπο που επιθυμούμε να οικοδομήσουμε την κοινωνία μας.

## 2.3 Φοροδιαφυγή

Όπως ήδη αναφέρθηκε, η έννοια της φορολογίας διαδραματίζει ζωτικό ρόλο στη σφαίρα της Δημοσίας Οικονομίας. Η μελέτη της φοροδιαφυγής είναι ένα πολυεπιστημονικό πεδίο που συνδυάζει πτυχές της κοινωνικής επιστήμης και του οικονομικού εγκλήματος.

Οι Allingham και Sandmo εισήγαγαν ένα θεωρητικό μοντέλο για τη φοροδιαφυγή, βασιζόμενοι στη μεθοδολογία της οικονομίας του εγκλήματος του Becker. Αυτό το μοντέλο, που θεωρείται το μέγιστο έργο τους, εξετάζει τη διαδικασία λήψης αποφάσεων των ορθολογικών φορολογουμένων ατόμων όσον αφορά τη συμμόρφωση με τους φορολογικούς κανονισμούς. Προσδιορίζει τρεις βασικούς παράγοντες που επηρεάζουν αυτή την απόφαση: τους φορολογικούς συντελεστές, τη σοβαρότητα των κυρώσεων και των προστίμων και την πιθανότητα ελέγχου. Ενώ αυτό το παραδοσιακό μοντέλο έχει κριτική για την απλότητά του, μεταγενέστερη έρευνα οδήγησε στην ανάπτυξη μεμονωμένων μοντέλων που στοχεύουν να συμπεριλάβουν τις διάφορες διαστάσεις της φοροδιαφυγής (Sandmo, 2005).

### 2.3.1 Ορισμός Φοροδιαφυγής

Σύμφωνα με το Λεξικό της Οξφόρδης, η φοροδιαφυγή χαρακτηρίζεται ως η παράνομη πράξη μη πληρωμής ή υποκαταβολής φόρων. Ομοίως, το Majesty's Revenue and Customs (HMRC), ο κυβερνητικός οργανισμός που είναι υπεύθυνος για τη συλλογή φόρων στο Ηνωμένο Βασίλειο, ορίζει τη φοροδιαφυγή ως τη σκόπιμη πράξη υποτίμησης των φορολογικών υποχρεώσεων κάποιου. Ο αντίκτυπος του μεγέθους της φοροδιαφυγής στα δημόσια οικονομικά έχει εκτεταμένες επιπτώσεις και διαδραματίζει κρίσιμο ρόλο στη συνολική οικονομική ανάπτυξη.

Επί του παρόντος, είναι επιτακτική ανάγκη να καθιερωθεί μια οριστική διαφοροποίηση μεταξύ της φοροδιαφυγής και τριών αλληλένδετων εννοιών: φοροαποφυγή, παραοικονομία και εγκληματικές δραστηριότητες. Αυτά τα όρια έχουν περιγραφεί συνοπτικά από τον Cowell (1990) και είναι κοινώς γνωστά ως όρια Cowell (Cowell, 1990).

### 2.3.2 Φοροδιαφυγή και Φοροαποφυγή

Η διάκριση μεταξύ φοροαποφυγής και φοροδιαφυγής έχει παρεξηγηθεί ευρέως, ιδιαίτερα μεταξύ ατόμων με περιορισμένη γνώση του θέματος. Ο Denis Healey, πρώην Υπουργός Οικονομικών της Βρετανίας, συνέκρινε περίφημα τη διαφορά μεταξύ των δύο ως τόσο ουσιαστική όσο ένας τοίχος φυλακής. Ωστόσο, μπορεί να υποστηριχθεί ότι οι όροι και οι ενέργειες που σχετίζονται με το καθένα δεν είναι πανομοιότυποι. Στην πραγματικότητα, ο διάσημος νομικός μελετητής Holmes (1916) έδωσε μια κλασική εξήγηση για την ανισότητα,



υποστηρίζοντας ότι η απόκλιση έγκειται κυρίως στις νομικές συνέπειες που συνοδεύουν κάθε πράξη, με τη φοροδιαφυγή να είναι αναμφισβήτητα ατιμώρητη.

Ο Sandmo (2005) υποστηρίζει ότι όταν ένας φορολογούμενος δεν αναφέρει φορολογητέο εισόδημα, εμπλέκεται σε μια παράνομη δραστηριότητα που τον καθιστά υπόλογο τόσο στις διοικητικές όσο και στις δικαστικές αρχές. Αντίθετα, η φοροαποφυγή λειτουργεί εντός των ορίων της φορολογικής νομοθεσίας και συνεπάγεται την εκμετάλλευση των κενών για την ελαχιστοποίηση των φορολογικών υποχρεώσεων (Sandmo, 2005).

Η βασική διάκριση έγκειται στην αντίληψη του φορολογούμενου για τον εντοπισμό και τις επακόλουθες επιπτώσεις που σχετίζονται με τη φοροδιαφυγή, ενώ η ενασχόληση με ενέργειες που συνιστούν φοροαποφυγή αμβλύνει αυτές τις ανησυχίες. Αξίζει να εξεταστεί η ειρωνική παρατήρηση του Slemrod ότι «οι φτωχοί φοροδιαφεύγουν ενώ οι πλούσιοι αποφεύγουν τους φόρους», υπογραμμίζοντας την ανάγκη διαφοροποίησης μεταξύ αυτών των δύο εννοιών (Slemrod and Yitzhaki, 2002). Αυτή η διαφοροποίηση θα πρέπει πρωτίτως να τονίζει τον παράνομο και τιμωρητικό χαρακτήρα της φοροδιαφυγής σε σύγκριση με τη φοροαποφυγή.

Σύμφωνα με το Αγγλικό Λεξικό της Οξφόρδης, η φοροδιαφυγή μπορεί να οριστεί ως η στρατηγική οργάνωση των οικονομικών θεμάτων κάποιου προκειμένου να εκπληρώσει τη νομική υποχρέωση να πληρώσει το λιγότερο δυνατό ποσό φόρου. Αυτό συνεπάγεται τη σκόπιμη διευθέτηση των οικονομικών θεμάτων κάποιου για να διασφαλιστεί ότι καταβάλλεται μόνο το ελάχιστο ποσό φόρου που απαιτείται από τη νομοθεσία.

Όταν εξετάζονται οι οικονομικές πτυχές και οι ενέργειες των ατόμων, αυτά τα περιστατικά μοιράζονται σημαντικές ομοιότητες, γεγονός που καθιστά δύσκολη τη διαφοροποίηση μεταξύ τους (Franzoni, 1999). Επιπλέον, από ηθική άποψη, το αποτέλεσμα είναι πανομοιότυπο, με αποτέλεσμα ορισμένοι να ισχυρίζονται ότι κατά την εξέταση της φοροδιαφυγής και της φοροαποφυγής, αυτές οι δύο έννοιες πρέπει να θεωρούνται συνώνυμες.

### 2.3.3 Φοροδιαφυγή και Παραοικονομία

Υπάρχουν δύο βασικές προοπτικές για τη συσχέτιση της φοροδιαφυγής και της παραοικονομίας. Η μία πλευρά πιστεύει ότι και τα δύο αυτά φαινόμενα συμβαδίζουν, πράγμα που σημαίνει ότι οι παράγοντες που καθορίζουν το ένα πρέπει να ισχύουν και για το άλλο. Από την άλλη πλευρά, ορισμένοι υποστηρίζουν ότι η παραοικονομία δεν περιλαμβάνει απλώς τη φοροδιαφυγή και πρέπει να αντιμετωπιστεί χωριστά. Επιπλέον, η παραοικονομία

περιλαμβάνει παράνομες δραστηριότητες που θα έπαυαν να υπάρχουν εάν ανακαλυφθούν κατά τη διάρκεια ενός ελέγχου, με αποτέλεσμα να μην δημιουργούνται φορολογικά έσοδα.

Ο Cowell υποστηρίζει ότι η ανάλυση οικονομικής συμπεριφοράς δεν αντιμετωπίζει περαιτέρω εμπόδια από τη συζήτηση για τη φοροδιαφυγή έναντι της φοροαποφυγής, όταν πρόκειται για την παραοικονομία. Η παραοικονομία δεν περιλαμβάνει απλώς τη φοροδιαφυγή. Περιλαμβάνει επίσης τη μη συμμόρφωση με τους κανόνες εργασίας, τα πρότυπα παραγωγής και άλλες μη φορολογικές νομικές απαιτήσεις (Cowell, 1985).

#### 2.3.4 Φοροδιαφυγή και Εγκληματικές Δραστηριότητες

Έχει διεξαχθεί εκτενής έρευνα σχετικά με τη συσχέτιση μεταξύ της φοροδιαφυγής και της οικονομίας των εγκληματικών δραστηριοτήτων. Αυτό θέτει το ερώτημα εάν η φοροδιαφυγή πρέπει να αντιμετωπίζεται ως μοναδικό υποσύνολο οικονομικών εγκλημάτων. Ο Cowell (1990) υποστηρίζει ότι οι φοροφυγάδες δικαιολογούν ξεχωριστή εκτίμηση ως λογικοί οικονομικοί εγκληματίες, για δύο βασικούς λόγους. Πρώτον, η φοροδιαφυγή συνιστά δόλια πράξη σε βάρος ενός εξαιρετικά εξειδικευμένου οικονομικού φορέα, της Κυβέρνησης. Η κυβέρνηση έχει την εξουσία να διαμορφώνει το οικονομικό τοπίο καθορίζοντας το πλαίσιο των οικονομικών σχέσεων, τη δομή των εν λόγω σχέσεων και τον τρόπο με τον οποίο φορολογούνται, συμπεριλαμβανομένου του καθορισμού φορολογικών συντελεστών. Επιπλέον, η κυβέρνηση είναι εξοπλισμένη με τους απαραίτητους μηχανισμούς για την παρακολούθηση και τον εντοπισμό παραβάσεων, καθώς και την εφαρμογή υποχρεωτικών κυρώσεων. Αντίθετα, σε περιπτώσεις άλλων εγκλημάτων όπως η κλοπή ή η απάτη, ούτε οι επιχειρήσεις ούτε τα άτομα διαθέτουν τα μέσα για να επιβάλουν την εξουσία τους ή να εξασφαλίσουν τη συμμόρφωση των οικονομικών οντοτήτων. Δεύτερον, η απόφαση συμμόρφωσης συχνά περιλαμβάνει την επιλογή αναφοράς της φοροδιαφυγής στις φορολογικές αρχές, χαρακτηριστικό που απουσιάζει σε άλλες μορφές εγκληματικής συμπεριφοράς (Cowell, 1990).

Συνοπτικά, η φοροδιαφυγή αναφέρεται στη σκόπιμη προσπάθεια που καταβάλλεται για απόκρυψη φορολογητέων κερδών και ελαχιστοποίηση των φορολογικών υποχρεώσεων προς τις κρατικές φορολογικές υπηρεσίες. Αυτή η ενέργεια διαφέρει από τη φοροαποφυγή, την παραοικονομία και διάφορες παράνομες επιχειρήσεις.

#### 2.4 Τρόποι μείωσης φοροδιαφυγής

Ο αποτελεσματικός σχεδιασμός ενός φορολογικού συστήματος από μόνος του δεν επαρκεί για τη μείωση της φοροδιαφυγής. Σε γενικές γραμμές, η κυβέρνηση μπορεί να μειώσει τη

φοροδιαφυγή παρέχοντας αγαθά που οι πολίτες επιθυμούν σε μεγαλύτερο βαθμό, πιο αποτελεσματικά ή καθιστώντας ακόμη πιο σαφές ότι οι φόροι είναι απαραίτητοι για την παροχή των κρατικών υπηρεσιών (Alm and Martinez-Vazquez, 2003).

Οι Allingham & Sandmo (1972) ανέπτυξαν ένα πρώτο μοντέλο ορθολογισμού μεγιστοποιώντας τη χρησιμότητα των φορολογουμένων που λαμβάνουν αποφάσεις υπό αβεβαιότητα. Κατέληξαν στο συμπέρασμα ότι η αύξηση των παραμέτρων αποτροπής, όπως για παράδειγμα η αύξηση της πιθανότητας ανίχνευσης και η αύξηση του ύψους των ποινών σε τέτοια περίπτωση, οδηγούν σε αύξηση των πραγματικά δηλωθέντων εισοδημάτων. Ωστόσο, επικρίθηκε η έρευνά τους καθώς τα συμπεράσματα στα οποία καταλήγει δε συνάδουν με τα εμπειρικά στοιχεία που αφορούν τη συμμόρφωση στη φορολογία εισοδήματος.

Οι πιο συχνοί τρόποι αντιμετώπισης της φοροδιαφυγής είναι η εντατικοποίηση των ελέγχων που διεξάγονται από τις φορολογικές αρχές και οι έρευνες που διεξάγουν οι ερευνητές πάνω σε μακροοικονομικά και μικροοικονομικά μεγέθη της οικονομίας και από τις οποίες μπορούν να προκύψουν ποσά υπολογισμού της φοροδιαφυγής (Slemrod & Weber, 2012). Και οι δύο τρόποι μπορούν να εξετάσουν ένα τυχαίο δείγμα ανθρώπων και να παράγουν μια εκτίμηση σχετικά με τη μη συμμόρφωση ανά άτομο. Ωστόσο, η εντατικοποίηση των ελέγχων συνεπάγεται μεγάλα κόστη για το κράτος.

Ένας τρόπος εύρεσης του πραγματικού εισοδήματος είναι η χρήση της ηλεκτρικής ενέργειας. Η χρήση της ηλεκτρικής ενέργειας από τις επιχειρήσεις αποτελεί ένδειξη του μεγέθους παραγωγής τους. Ακόμη, τα χρήματα (μετρητά) αποτελούν ένδειξη φοροδιαφυγής. Ο Feige (1990) εκτιμά το μέγεθος της παραοικονομίας υποθέτοντας ότι το μεγαλύτερο μέρος της μη δηλωμένης δραστηριότητας πραγματοποιείται με μετρητά, ενώ υπάρχει και ένα έτος «βάσης» κατά το οποίο δεν υπήρχε φοροδιαφυγή. Η πιο σύνθετη εμπειρική προσέγγιση για τη μέτρηση της παραοικονομίας και της φοροδιαφυγής χρησιμοποιεί πληροφορίες για τα ίχνη του πραγματικού εισοδήματος, τα ίχνη της μη συμμόρφωσης και εκτιμήσεις. Αυτή η προσέγγιση είναι γνωστή ως MIMIC (multiple indicators – multiple causes), καθώς λαμβάνει υπόψη πολλαπλούς δείκτες και πολλαπλές αιτίες, αναπτύχθηκε από τον Zellner (1970) και εφαρμόστηκε πρώτη φορά από τους Frey & Weck-Hanneman (1984).

Η ικανότητα του κράτους να επιβάλλει αποτελεσματικά την είσπραξη των φόρων αποτελεί ένα από τα θεμελιώδη στοιχεία της κυβερνητικής ικανότητας και αποτελεί ως εκ τούτου κινητήρια δύναμη οικονομικής ανάπτυξης. Η φοροδιαφυγή δημιουργεί σημαντικές απώλειες και στρεβλώσεις στα δημόσια έσοδα (Slemrod and Yitzhaki, 2002). Σύμφωνα με έρευνα που

διεξήχθη με πρωτοβουλία της Τράπεζας της Ιταλίας διαπιστώθηκε ότι το 76% των συμμετεχόντων πιστεύει ότι η καταπολέμηση της φοροδιαφυγής πρέπει να συγκαταλέγεται στις άμεσες προτεραιότητες της ιταλικής δημόσιας πολιτικής.

Ο Bergman (2009) καταλήγει στο συμπέρασμα ότι τα μέτρια αποτελέσματα στα κράτη που τηρούν το νόμο παράγουν καλύτερα αποτελέσματα από ό,τι μια καλή διοίκηση σε ένα κράτος που χαρακτηρίζεται από τη φοροδιαφυγή. Η αντιμετώπιση των παράνομων κεφαλαιακών ροών δεν απαιτεί απλώς αυστηρότερη εφαρμογή της φορολογικής πολιτικής και της φορολογικής διοίκησης, αλλά μετατροπή της στάσης των πολιτών απέναντι στους πολιτικούς θεσμούς που διαμορφώνουν το πλαίσιο της διακυβέρνησης.

Οι Bazart & Pickhardt (2009) εξέτασαν την επίδραση θετικών ανταμοιβών στην συμπεριφορά φοροδιαφυγής με τη χρήση ατομικών κερδών από λαχειοφόρες αγορές για όσους συμμορφώνονται στις φορολογικές υποχρεώσεις. Διαπίστωσαν ότι οι θετικές ανταμοιβές οδηγούν σε σημαντικά υψηλότερο ποσοστό φορολογικής συμμόρφωσης. Ακόμη, διαπίστωσαν ότι οι άντρες φοροδιαφεύγουν σε μεγαλύτερο βαθμό από τις γυναίκες. Τα κίνητρα των κρατικών υπαλλήλων για την επιβολή είσπραξης των φόρων αποτελούν εξίσου καθοριστικό παράγοντα του επιπέδου συμμόρφωσης.

Η Christians (2014) υποστήριξε ότι η ηθική μπορεί να λειτουργήσει ως ένα μέτρο καταπολέμησης της φοροδιαφυγής. Υποστηρίζει την εναλλακτική πρόταση η συμπεριφορά των φορολογουμένων να διέπεται από το νόμο και όχι από την κοινωνική κύρωση και κατακραυγή. Αυτή η μέθοδος ισχυρίζεται ότι έχει καλύτερες πιθανότητες να οδηγήσει τη φορολογική πολιτική σε μεγαλύτερη συνοχή μακροπρόθεσμα. Αυτή η άποψη φυσικά δε σημαίνει ότι το κοινό δεν πρέπει να συμμετέχει στον πολιτικό λόγο. Αντιθέτως, το κοινό είναι το πιο κατάλληλο να ζητά απόλυτη διαφάνεια στη διακυβέρνηση λειτουργώντας ως ελεγκτικός μηχανισμός και διευθετώντας τα προβλήματα φορολογικής πολιτικής.

Οι Casaburi & Troiano (2016) μελέτησαν την εφαρμογή του προγράμματος Ghost Buildings που αφορούσε μια πολιτική καταπολέμησης της φοροδιαφυγής στην Ιταλία με καινοτόμες τεχνολογίες παρακολούθησης ώστε να εντοπιστούν κτήρια κρυμμένα από τις φορολογικές αρχές. Το πρόγραμμα προσκόμισε οφέλη τόσο οικονομικά όσο και μη για όσους δεν φοροδιαφεύγουν και μια αύξηση στις δαπάνες της τοπικής αυτοδιοίκησης. Κατέληξαν στο συμπέρασμα ότι οι σύγχρονες παρεμβάσεις που βελτιώνουν την τεχνολογία της επιβολής φόρων όπως οι αναφορές τρίτων, η διασταύρωση στοιχείων ή καλύτεροι αλγόριθμοι ελέγχου,

μπορούν να μειώσουν σημαντικά τη φοροδιαφυγή. Απαιτούνται όμως και πολιτικά κίνητρα υιοθέτησης τέτοιων μεθόδων εξάλειψης της φοροδιαφυγής.

Συνήθως, η φοροδιαφυγή αντιμετωπίζεται με αρνητικές κυρώσεις όπως τα πρόστιμα ή ακόμη και η φυλάκιση για όσους φοροδιαφεύγουν παρανόμως. Για αυτό και πολλές φορές η επιτυχία τέτοιων μέτρων είναι αμφισβητήσιμη (Cristie & Holzner, 2006).

## Κεφάλαιο 3ο: Μηχανική Μάθηση και Φοροδιαφυγή

Στις μέρες μας, οι έννοιες της μηχανικής μάθησης και της φοροδιαφυγής είναι ευρέως διαδεδομένες και η χρήση τους υποβοηθά τους εφοριακούς και τους λογιστές στις καθημερινές τους πρακτικές. Η φοροδιαφυγή είναι ένα πολύ σοβαρό οικονομικό έγκλημα και η μηχανική μάθηση μπορεί να αξιοποιηθεί για την ανίχνευση και την πρόληψη αυτών των ανεπιθύμητων περιπτώσεων. Η σύνδεση της μηχανικής μάθησης με τη φοροδιαφυγή μπορεί να γίνει μέσω διαφορών τρόπων, ανάλογα με την φύση του προβλήματος και του αλγορίθμου που χρησιμοποιήθηκε.

Στη συνέχεια του κεφαλαίου παρουσιάζεται η κυριότερη πρόσφατη αρθρογραφία σχετικά με την σύνδεση των δύο εννοιών και έπειτα αναλύεται ο ρόλος των τεχνικών μηχανικής μάθησης στον εντοπισμό και την πρόληψη της φοροδιαφυγής όπως αυτή προκύπτει από την ανάλυση της διαθέσιμης βιβλιογραφίας. Τέλος, θα αναφερθούν εφαρμογές της μηχανικής μάθησης στη λογιστική διαχείριση αλλά και στη φοροδιαφυγή.

### 3.1 Πρόσφατη αρθρογραφία

Σήμερα υπάρχουν αρκετές έρευνες στη βιβλιογραφία ως προς τη διεργασία αξιοποίησης και εφαρμογής τεχνικών μηχανικής μάθησης στη λογιστική. Οι επιδιώξεις και τα ευρήματα μερικών συγγραφέων περιγράφονται στην επόμενη παράγραφο.

Οι **N. Visitpanya και T. Samanchuen** στο άρθρο τους «A Machine Learning Approach to Identifying Suspicious Tax Evasion Behavior in Public Financial Data» που δημοσιεύτηκε το 2023, αναφέρουν ότι ο εντοπισμός της «ύποπτης» συμπεριφοράς που σχετίζεται με φοροδιαφυγή απαιτεί εξελιγμένες τεχνικές και σύγχρονες μεθόδους επεξεργασίας (Visitpanya and Samanchuen, 2023).

Πιο αναλυτικά, η μελέτη αυτή παρουσιάζει μια μέθοδο για τον εντοπισμό της προδιάθεσης ενός ατόμου να φοροδιαφύγει χρησιμοποιώντας οικονομικά δεδομένα, που είναι διαθέσιμα δημοσίως. Τα δεδομένα αυτά, λήφθηκαν από δημόσια πηγή και εφαρμόστηκε μηχανική μάθηση για τον εντοπισμό της «ύποπτης» αυτής συμπεριφοράς με βάση κάποια κριτήρια. Η ακρίβεια των αποτελεσμάτων τους είναι 95% χρησιμοποιώντας εποπτευόμενα μοντέλα. Περισσότερα από 480 δισεκατομμύρια δολάρια χανόντουσαν λόγω της φοροδιαφυγής των εταιρειών στον κόσμο. Για αυτόν τον λόγο, επιλεκτικέ μια μεγάλη εταιρεία στην Ταϊλάνδη που τα δεδομένα είναι διαθέσιμα στο κοινό. Αφού εκφωνήθηκαν οι πηγές δεδομένων,

μελετήθηκαν τα κριτήρια αξιολόγησης κινδύνου. Με τα δεδομένα και τα κριτήρια που αποκτήθηκαν, χρησιμοποιήθηκε η μηχανική μάθηση για την ανίχνευση μιας τέτοιας συμπεριφοράς. Μετά την ολοκλήρωση αυτής της μελέτης, προτείνουν να εφαρμοστεί και με αλγόριθμο μάθησης χωρίς επίβλεψη, την επικύρωση δεδομένων που συλλέχθηκε παρόμοιο με εκείνα της κυβέρνησης και συλλέγεται από εισηγμένες εταιρίες στο Stock Exchange of Thailand. Προβλήματα που αναφέρουν είναι ότι δεν γνωρίζουν εάν αντιπροσωπεύει μικρομεσαίες επιχειρήσεις που δεν εγγράφονται στο Stock Exchange of Thailand, παρουσιάζονται μόνο τα γενικά χαρακτηριστικά, δεν δείχνει διαφορές μεταξύ των επιμέρους επιχειρηματικών τομέων, λόγω του περιορισμού του όγκου των δεδομένων και τέλος ενδέχεται να μην είναι ακόμη διαθέσιμο ένα σύνολο δοκιμών για την αξιολόγηση των επιδόσεων των μοντέλων. Στο σημείο αυτό αξίζει να σημειωθεί η παρατήρηση των αρθρογράφων ότι η προστασία των ευαίσθητων φορολογικών δεδομένων αποτελεί πρόκληση για τους ερευνητές που ασχολούνται με την επεξεργασία τους και επιδιώκουν να τα χρησιμοποιήσουν ελεύθερα, ιδίως δεδομένα από τον τομέα των οικονομικών και της φορολογίας. Επομένως, είναι σημαντική η γνώση των διατάξεων για την προστασία των δεδομένων και τη σημασία της διατήρησης της ιδιωτικότητας των φορολογούμενων κατά τη διαχείριση αυτών των δεδομένων.

Στη συνέχεια, οι **Belle Fille Murorunkwere, Dominique Haughton, Joseph Nzabanita, Francis Kipkogei και Ignace Kabano** στο άρθρο τους «Predicting tax fraud using supervised machine learning approach» που δημοσιεύτηκε το 2023, στοχεύουν στην πρόβλεψη χαρακτηριστικών της φορολογικής απάτης χρησιμοποιώντας ένα ισχυρό μοντέλο εποπτευόμενης μηχανικής μάθησης (Murorunkwere et al., 2023). Αυτή η μελέτη αξιολογεί την αποτελεσματικότητα πολλαπλών εποπτευόμενων μοντέλων μηχανικής μάθησης για τον εντοπισμό παραγόντων φορολογικής απάτης σε ένα ακριβές σύνολο δεδομένων με πολλούς τύπους φόρων.

Πιο αναλυτικά, παρουσιάζεται ένα πλαίσιο όπου ένας εμπειρογνώμονας απάτης μπορεί να χρησιμοποιήσει ένα μοντέλο μηχανικής μάθησης και ένα εφαρμοσμένο μοντέλο που του προσφέρει άμεση ανατροφοδότηση χρησιμοποιώντας εποπτευόμενα μοντέλα μηχανικής μάθησης. Μια διαδεδομένη μορφή φορολογικής απάτης είναι η ελλιπής δήλωση, που περιλαμβάνει την υποβολή ενός εντύπου φορολογικής δήλωσης με μειωμένη φορολογική βάση. Η παραβίαση των φορολογικών κανονισμών στις Ηνωμένες Πολιτείες οδήγησε σε εκτιμώμενη απώλεια περίπου 450 δισεκατομμυρίων δολαρίων που προκλήθηκε από μεμονωμένους μη συμμορφούμενους φορολογούμενους. Για να συγκαλύψουν και να

μειώσουν την πραγματική φορολογική ευθύνη των μεμονωμένων ενδιαφερομένων μερών τους, τα καταφύγια συνδυάζουν διάφορες επιχειρηματικές δομές, όπως εταιρείες, συμπράξεις και καταπιστεύματα, σε πολύπλοκα δίκτυα συναλλαγών. Σύμφωνα με την έκθεση του Διεθνούς Νομισματικού Ταμείου (ΔΝΤ), η φοροδιαφυγή μπορεί να κοστίζει στα αναπτυσσόμενα έθνη έως και 213 δισεκατομμύρια δολάρια ετησίως. Οι φορολογικές αρχές χρησιμοποιούν παραδοσιακά δύο τρόπους για την καταπολέμηση της φορολογικής απάτης: την εμπειρία των ελεγκτών και τα συστήματα που βασίζονται σε κανόνες. Η πρώτη μέθοδος περιλαμβάνει την τυχαιοποίηση των φορολογικών δηλώσεων και τον έλεγχό τους με τη χρήση της διαίσθησης, της εμπειρίας και της γνώσης του τομέα των φορολογικών ελεγκτών. Σήμερα, η ανίχνευση της φορολογικής απάτης με τη χρήση προσεγγίσεων μηχανικής μάθησης έχει γίνει ένας πολύ ενεργός τομέας ενδιαφέροντος στην ερευνητική κοινότητα. Για την ανίχνευση περιπτώσεων απάτης, οι ειδικοί και οι ερευνητές στον τομέα της ανίχνευσης απάτης έχουν χρησιμοποιήσει μια σειρά μεθόδων, όπως η μηχανική μάθηση. Το μοντέλο των νευρωνικών δικτύων επέτρεψε την τμηματοποίηση των φορολογουμένων και προσέφερε μια απλή εκτίμηση της πιθανότητας να προσπαθήσει ένας μεμονωμένος φορολογούμενος να αποφύγει τους φόρους. Ο αλγόριθμός τους ήταν 84,3% ακριβής στον εντοπισμό όσων προσπαθούν να αποφύγουν την καταβολή φόρων, γεγονός που υποδηλώνει ότι μπορεί να εφαρμοστεί και σε άλλες μορφές φόρων. Για να δημιουργήσουν ένα θεωρητικό και μεθοδολογικό πλαίσιο για τις εφαρμογές της ανάλυσης δεδομένων στους φόρους, οι ερευνητές χρησιμοποίησαν μηχανική μάθηση και προηγμένη ανάλυση δεδομένων για τον εντοπισμό της φορολογικής απάτης. Απέδειξαν ότι οι φόροι είναι ένας από τους ιδανικούς τομείς για εφαρμογές μηχανικής μάθησης με επίβλεψη. Πρότειναν να χρησιμοποιηθούν διάφορες προσεγγίσεις μηχανικής μάθησης σε μεγάλης κλίμακας πραγματικά σύνολα δεδομένων από φορολογικές αρχές, γεγονός που θα βοηθούσε στην ανακάλυψη συγκεκριμένων αλγορίθμων μηχανικής μάθησης που αποδίδουν καλά κατά την ανίχνευση μιας συγκεκριμένης τύπου απάτης.

Τα ευρήματα αποκαλύπτουν ότι η φοροδιαφυγή σχετίζεται άμεσα με το χρόνο λειτουργίας μιας επιχείρησης, δηλαδή από τη χρονική απόσταση από την έναρξή της μέχρι τη στιγμή που ελέγχθηκε, και το γεγονός ότι είναι πιο πιθανό να φοροδιαφύγουν οι εγχώριες επιχειρήσεις, οι φορολογούμενοι που ασχολούνται με το εμπόριο (εισαγωγές – εξαγωγές) και οι φορολογούμενοι που έχουν διαπράξει στο παρελθόν το αδίκημα αυτό χωρίς τις αντίστοιχες επιπτώσεις. Από τη μελέτη αυτή προκύπτουν στοιχεία που αποτελούν πολύτιμο εργαλείο για τους ελεγκτές του και για τους δημόσιους φορείς ως προς τη φορολογική πολιτική.



Έπειτα, οι **Vardan Baghdasaryan, Hrant Davtyan, Arsine Sarikyan & Zaruhi Navasardyan** στο άρθρο τους «Improving Tax Audit Efficiency Using Machine Learning: The Role of Taxpayer's Network Data in Fraud Detection» που δημοσιεύθηκε το 2022 αναφέρουν πως χρησιμοποιώντας το σύνολο των Αρμένιων φορολογουμένων επιχειρήσεων που λειτουργούν υπό ένα τυπικό φορολογικό καθεστώς, αναπτύσσουν ένα μοντέλο πρόβλεψης φορολογικής απάτης που βασίζεται σε εργαλεία μηχανικής μάθησης (Baghdasaryan et al., 2022).

Η διαδικασία συλλογής δεδομένων παρουσίασε αρκετές δυσκολίες. Η ετερογένεια και ο μικρός αριθμός του δείγματος, καθώς και το γεγονός πως οι επιχειρηματίες συχνά φοροδιαφεύγουν αποτελούν τα κυριότερα προβλήματα που αντιμετώπισαν. Τελικά συγκέντρωσαν όλα τα απαραίτητα δεδομένα μέσω των φορολογικών δηλώσεων και κάποιες πρόσθετες πληροφορίες. Ανάμεσα στους πιο σημαντικούς παράγοντες πρόβλεψης φορολογικής απάτης, λαμβάνονται το παρελθόν που σχετίζεται με την απάτη αυτή και τους σχετικούς ελέγχους, το διοικητικό κόστος και η οικονομική δραστηριότητα στο εξωτερικό. Σύμφωνα με το αποτέλεσμα τους, το ποσοστό απάτης στην κορυφή είναι περίπου 1,85 φορές υψηλότερο σε σύγκριση με το μέσο παρατηρούμενο. Το πλαίσιο εξαγωγής συμπερασμάτων επιτρέπει να λάβουν εκτιμήσεις της σχετικής σημασίας των χαρακτηριστικών του μοντέλου. Οι πληροφορίες αυτές μπορούν να χρησιμοποιηθούν από το τμήμα διαχείρισης κινδύνων για την αναβάθμιση των κανόνων που εφαρμόζονται.

Τα δύο κυριότερα ευρήματα, τα οποία είναι δυνατό να γενικευτούν και να χρησιμοποιηθούν από τις διάφορες ελεγκτικές αρχές είναι:

- Πρώτον, εστιάζοντας στη βαθμολογία ανύψωσης της κορυφαίας δεκαετίας, αποδεικνύεται ότι ακόμη και μοντέλα μέτριας ακρίβειας μπορούν να βελτιώσουν την υπάρχουσα ακρίβεια των προσεγγίσεων που βασίζονται σε κανόνες.
- Δεύτερον, και σημαντικότερο, αποδεικνύεται ότι οι πληροφορίες που περιέχονται από το δίκτυο προμηθευτών και αγοραστών της προς εξέταση επιχείρησης μπορούν να χρησιμοποιηθούν σε περίπτωση που δεν είναι διαθέσιμοι οι σημαντικοί παράγοντες πρόβλεψης φορολογικής απάτης που προαναφέρθηκαν. Αυτό είναι ένα ιδιαίτερα σημαντικό εργαλείο σε διάφορες περιπτώσεις, όπως για παράδειγμα στις νεοϊδρυθείσες εταιρείες, οι οποίες σε αντίθετη περίπτωση θα ήταν υποβαθμισμένες ως προς την πιθανότητα απάτης.

Οι **Lahann, Scheid και Fettke** στο άρθρο «Utilizing Machine Learning Techniques to Reveal VAT Compliance Violations in Accounting Data» που δημοσιεύτηκε το 2019, μελέτησαν τις τεχνικές μηχανικής μάθησης προκειμένου να αποδείξουν τον τρόπο που δύναται να αξιοποιηθούν για να μειώσουν κατά πολύ τους κινδύνους και να αυξήσουν το ενδεχόμενο να εναρμονιστούν σύμφωνα με το υπόδειγμα του ΦΠΑ (Lahann, Scheid and Fettke, 2019).

Πιο συγκεκριμένα, εξέτασαν μια ευρέως εφαρμοζόμενη προσέγγιση για την επαλήθευση των κανονισμών ΦΠΑ στο πλαίσιο των συστημάτων διαχείρισης επιχειρησιακών πόρων (ERP) και τόνισαν τις αδυναμίες και τις ευαισθησίες σε σφάλματα. Επιπλέον, έδειξαν ότι οι τεχνικές μηχανικής μάθησης μπορούν να αξιοποιηθούν για την ελαχιστοποίηση των κινδύνων και την αύξηση της συμμόρφωσης με τον ΦΠΑ. Χρησιμοποίησαν έναν ταξινομητή επιβλεπόμενης μάθησης για την πρόβλεψη των φορολογικών υποκειμένων και των αντίστοιχων φορολογικών συντελεστών με βάση τις σχετικές πληροφορίες παραστατικών των ημερολογιακών αναφορών. Συγκρίνοντας τα αποτελέσματα του μοντέλου με το υπάρχον σύστημα που βασίζεται σε κανόνες ενός συστήματος ERP, ανακάλυψαν πιθανές ανωμαλίες και ζητήματα συμμόρφωσης. Τα αποτελέσματα επικυρώθηκαν από εμπειρογνώμονες ΦΠΑ της εταιρείας.

Στην έρευνά τους, το καθολικό σύνολο δεδομένων μιας γνωστής και ισχυρής εταιρείας στη χημική βιομηχανία συλλέχθηκε από το σύστημα ERP της εταιρείας. Τα ευρήματα φάνέρωσαν ότι οι αλγόριθμοι μηχανικής μάθησης είναι σε θέση να βρουν σφάλματα που θα συνέτειναν σε παραβιάσεις που είναι αντίθετες προς την εναρμόνιση με το κράτος, γεγονός που γεννά ελπίδα για την συμμόρφωση με τις αρχές και τους κανόνες φορολογικής εναρμόνισης στα συστήματα πληροφορικής.

### 3.2 Ο ρόλος των τεχνικών μηχανικής μάθησης στον εντοπισμό και την πρόληψη της φοροδιαφυγής

Οι αλγόριθμοι μηχανικής μάθησης έχουν αποδειχθεί χρήσιμοι στην ανάλυση οικονομικών δεδομένων για τον εντοπισμό προτύπων φοροδιαφυγής. Μπορούν να αναλυθούν προηγούμενα πρότυπα συμπεριφοράς για τον εντοπισμό πιθανής φορολογικής απάτης και οι αλγόριθμοι μπορούν να εκτιμήσουν τον κίνδυνο φοροδιαφυγής. Με τη χρήση καταγγελιών πολιτών, προηγμένων δυνατοτήτων συλλογής και ανάλυσης δεδομένων και αυτοματοποίησης, η διαδικασία εντοπισμού περιπτώσεων φοροδιαφυγής μπορεί να επιταχυνθεί και να γίνει πιο αποτελεσματική (Χατζηνικολάου, 2023).

Οι αλγόριθμοι μηχανικής μάθησης χρησιμοποιούν στατιστικά μοντέλα για τον εντοπισμό προτύπων σε μεγάλα σύνολα δεδομένων. Αυτοί οι αλγόριθμοι μπορούν να εκπαιδευτούν για τον εντοπισμό ανωμαλιών στις φορολογικές δηλώσεις, τον εντοπισμό πιθανών φοροφυγάδων και την πρόβλεψη της πιθανότητας μη συμμόρφωσης. Ένα από τα σημαντικά οφέλη της χρήσης μηχανικής μάθησης στον εντοπισμό φοροδιαφυγής είναι η ικανότητά του να επεξεργάζεται τεράστιες ποσότητες δεδομένων γρήγορα και με ακρίβεια. Για παράδειγμα, η Υπηρεσία Εσωτερικών Εσόδων (IRS) χρησιμοποιεί αλγόριθμους μηχανικής μάθησης για την ανάλυση των φορολογικών δηλώσεων και τον εντοπισμό ύποπτων μοτίβων που μπορεί να υποδηλώνουν φοροδιαφυγή. Ένα άλλο παράδειγμα επιτυχημένης εφαρμογής μηχανικής εκμάθησης στον εντοπισμό φοροδιαφυγής είναι η χρήση αλγορίθμων ανίχνευσης ανωμαλιών από το Αυστραλιανό Φορολογικό Γραφείο. Το Αυστραλιανό Φορολογικό Γραφείο χρησιμοποιεί αλγόριθμους μηχανικής μάθησης για να αναλύσει τις φορολογικές δηλώσεις και να εντοπίσει ασυνήθιστη συμπεριφορά που μπορεί να υποδηλώνει μη συμμόρφωση. Αυτή η προσέγγιση βοήθησε το Αυστραλιανό Φορολογικό Γραφείο να εντοπίσει δόλιες δραστηριότητες, με αποτέλεσμα την αυξημένη φορολογική συμμόρφωση.

Οι τεχνικές μηχανικής μάθησης χρησιμοποιούνται όλο και περισσότερο για τον εντοπισμό και την πρόληψη της φοροδιαφυγής και προσφέρουν πολλά πλεονεκτήματα και περιορισμούς. Ένα από τα πλεονεκτήματα της χρήσης μηχανικής μάθησης είναι ότι μπορεί να επεξεργαστεί τεράστιες ποσότητες δεδομένων γρήγορα και με ακρίβεια, κάτι που μπορεί να βοηθήσει στον εντοπισμό προτύπων και ανωμαλιών που μπορεί να υποδηλώνουν φοροδιαφυγή. Για την καταπολέμηση της φοροδιαφυγής, οι φορολογικές αρχές χρησιμοποιούν τεχνικές μηχανικής μάθησης. Για αυτό το λόγο, αναπτύσσεται ένα σύστημα μηχανικής μάθησης που ενσωματώνει διάφορες τεχνικές για την επεξεργασία δεδομένων και την εξαγωγή πληροφοριών (Παλαιτσάκης, 2023).

Τα συστήματα μηχανικής μάθησης που βασίζονται σε τεχνικές μάθησης με επίβλεψη έχουν δείξει πολλά υποσχόμενη βελτίωση της αποτελεσματικότητας και της ακρίβειας των φορολογικών ελέγχων και των διαδικασιών συμμόρφωσης. Για παράδειγμα, μια μελέτη που διεξήχθη στην Ιταλία διαπίστωσε ότι η αντικατάσταση των ελέγχων, με φορολογούμενους που επιλέχθηκαν από εκπαιδευμένο αλγόριθμο, αύξησε την ανιχνευμένη φοροδιαφυγή έως και 38% και μείωσε το ποσό της φοροδιαφυγής κατά 29% (Eur-Lex, 2022).

Τα συστήματα μηχανικής μάθησης που βασίζονται σε τεχνικές μάθησης χωρίς επίβλεψη ή περιγραφικές αναλύσεις είναι ιδιαίτερα χρήσιμα για τον εντοπισμό ύποπτων συμπεριφορών.

Επιπλέον, τέτοια συστήματα μπορούν να εκπαιδευτούν ώστε να μαθαίνουν από προηγούμενες υποθέσεις φοροδιαφυγής, κάτι που μπορεί να βοηθήσει στον εντοπισμό νέων πιθανών υποθέσεων.

Η σημασία της δίκαιης φορολογίας και η καταπολέμηση της φορολογικής απάτης και φοροδιαφυγής είναι πρωταρχικής σημασίας. Οι φορολογικές αρχές αναγνωρίζουν τη σημασία των τεχνικών μηχανικής μάθησης για τον εντοπισμό και την πρόληψη της φοροδιαφυγής και χρησιμοποιούν όλο και περισσότερο αυτές τις τεχνολογίες για να προστατεύονται από δόλιες συμπεριφορές (Eur-Lex, 2022).

Ωστόσο, υπάρχουν επίσης ορισμένοι περιορισμοί στη χρήση τεχνικών μηχανικής μάθησης. Για παράδειγμα, η ακρίβεια των αποτελεσμάτων εξαρτάται από την ποιότητα και την πληρότητα των διαθέσιμων δεδομένων. Οι αλγόριθμοι μηχανικής μάθησης είναι τόσο καλοί όσο και τα δεδομένα στα οποία εκπαιδεύονται, επομένως εάν τα δεδομένα εκπαίδευσης είναι μεροληπτικά ή ελλιπή, το σύστημα μπορεί να δημιουργήσει ανακριβή αποτελέσματα.

Η χρήση της μηχανικής μάθησης για τον εντοπισμό και την πρόληψη της φοροδιαφυγής εγείρει ηθικές ανησυχίες. Μία από τις κύριες ανησυχίες είναι η πιθανότητα μεροληψίας στους αλγόριθμους μηχανικής μάθησης. Αυτοί οι αλγόριθμοι βασίζονται σε ιστορικά δεδομένα για να μάθουν και εάν τα δεδομένα είναι μεροληπτικά, οι αλγόριθμοι μπορεί να παράγουν μεροληπτικά αποτελέσματα. Αυτό θα μπορούσε να οδηγήσει σε άδικη μεταχείριση ορισμένων ομάδων φορολογουμένων, επιδεινώνοντας περαιτέρω τις υπάρχουσες ανισότητες. Μια άλλη ηθική ανησυχία είναι η έλλειψη διαφάνειας και υπευθυνότητας στη χρήση των αλγορίθμων μηχανικής μάθησης. Οι φορολογούμενοι μπορεί να μην γνωρίζουν πώς χρησιμοποιούνται τα δεδομένα τους ή πώς λαμβάνουν αποφάσεις οι αλγόριθμοι. Αυτή η έλλειψη διαφάνειας θα μπορούσε να διαβρώσει την εμπιστοσύνη στις φορολογικές αρχές και να υπονομεύσει τις προσπάθειες συμμόρφωσης.

Εν κατακλείδι, η σύνδεση μηχανικής μάθησης με τη φοροδιαφυγή μπορεί να γίνει μέσω διαφόρων τρόπων, ανάλογα με τη φύση του προβλήματος και τον αλγόριθμο μηχανικής μάθησης που θα χρησιμοποιηθεί. Έτσι, με τα σωστά δεδομένα και αλγόριθμους, οι τεχνικές μηχανικής μάθησης μπορούν να αποτελέσουν πολύτιμο εργαλείο για τον εντοπισμό και την πρόληψη της φοροδιαφυγής.

### 3.3 Μηχανική μάθηση: εφαρμογή στη λογιστική διαχείριση

Η μηχανική μάθηση έχει ανοίξει μια νέα εποχή στη λογιστική βελτιώνοντας την ακρίβεια στην ανάλυση των λογιστικών δεδομένων. Με αυτή την τεχνολογία, οι επιχειρήσεις μπορούν να έχουν πιο αξιόπιστα και ακριβή αποτελέσματα, επιτρέποντάς τους να λαμβάνουν πιο ενημερωμένες στρατηγικές αποφάσεις στον οικονομικό τομέα (Weil, De Silva and Ward, 2014).

Η σωστή χρήση και αξιοποίηση των αλγορίθμων μηχανικής μάθησης δύναται να προσφέρουν στους λογιστές καλύτερη ανάλυση δεδομένων που είναι σταθερή και ουσιαστική. Τα εργαλεία μηχανικής μάθησης τα οποία δημιουργούνται από εξειδικευμένους επιστήμονες σε αυτό το αντικείμενο δύναται να αποτελούν μια σειρά από εργασίες που διευκολύνουν το έργο των λογιστών. Λόγου χάρη, αντί για δεδομένα δειγματοληψίας, το καθολικό πλαίσιο μιας επιχείρησης μπορεί να ελεγχθεί με αυτοματοποίηση.

Οι αλγόριθμοι μηχανικής μάθησης έχουν την ικανότητα να τσεκάρουν και να μελετούν τα δεδομένα, να εντοπίζουν προβλήματα και λάθη και να δημιουργούν μια λίστα με ακραίες τιμές για επιθεώρηση από τους ελεγκτές. Αντί να καταναλώνουν όλο τον διαθέσιμο χρόνο τους στο τσεκάρισμα των στοιχείων, οι ελεγκτές δύναται να αξιοποιήσουν τις ικανότητές τους στη μελέτη και την διαπίστωση της αιτίας πίσω από ένα πρότυπο ή ανωμαλία. Ακόμη και αν αποφασίσουν να δουλέψουν από κοινού με ιδιοκτήτες μικρών επιχειρήσεων που δεν κατέχουν την λογιστική επιστήμη, και πάλι οι ελεγκτές είναι υποχρεωμένοι να προβούν σε ενημέρωση των βιβλίων των εκάστοτε εταιρειών, γεγονός που συνεπάγεται ότι θα αφιερώσουν πολλή ώρα προκειμένου να ανακαλούν στη μνήμη τους πού να κατανέμουν τις συναλλαγές τους. Επιπλέον, σε όλη αυτή τη διεργασία ενδέχεται οι λογαριασμοί να μην είναι εντελώς ορθοί και ακριβείς και να υπάρχουν κενά προς την οικονομική τους απόδοση. Παράλληλα με την εξοικονόμηση χρόνου, εάν οι ελεγκτές δύναται να επιθεωρούν κάθε πτυχή συναλλαγής μιας επιχείρησης, ενώ τα οικονομικά τους στοιχεία θα είναι πιο ακριβή και οι ελεγκτές θα είναι σε θέση να ασχολούνται πιο διεξοδικά στην μελέτη των οικονομικών δεδομένων για να προσφέρουν συμφέρουσες συμβουλές στους πελάτες τους.

Η μηχανική μάθηση έχει πολλαπλές εφαρμογές στον τομέα της λογιστικής. Μερικά από τα πιο σημαντικά είναι (Üzoğlu, 2020):

- Ανάλυση κινδύνου: Η μηχανική μάθηση μπορεί να βοηθήσει στον εντοπισμό μοτίβων και τάσεων στα λογιστικά δεδομένα που υποδεικνύουν πιθανούς χρηματοοικονομικούς κινδύνους. Αυτό επιτρέπει στις επιχειρήσεις να λαμβάνουν προληπτικά μέτρα και να μειώνουν την πιθανότητα ζημιών.

- Ανίχνευση απάτης: Αναλύοντας μεγάλους όγκους δεδομένων, οι αλγόριθμοι μηχανικής μάθησης μπορούν να εντοπίσουν ανωμαλίες και ύποπτα μοτίβα που θα μπορούσαν να υποδηλώνουν την παρουσία λογιστικής απάτης. Αυτό βοηθά τις επιχειρήσεις να προλαμβάνουν και να εντοπίζουν την απάτη πιο αποτελεσματικά (Cho et al., 2020). Ακόμα τις επιτρέπει να αναλαμβάνουν γρήγορη δράση για την πρόληψη ή τη διερεύνηση πιθανής απάτης, η οποία με τη σειρά της συμβάλλει στην προστασία των περιουσιακών τους στοιχείων και στη βελτίωση της ακεραιότητας των λογιστικών δεδομένων.
- Βελτιστοποίηση διαδικασιών: Η μηχανική μάθηση μπορεί να βοηθήσει στην αυτοματοποίηση εργασιών λογιστικής ρουτίνας, όπως η συμφωνία λογαριασμών ή η ταξινόμηση συναλλαγών. Μία από τις πιο κοινές χρήσεις της μηχανικής μάθησης στη λογιστική είναι στην ταξινόμηση και κατηγοριοποίηση των χρηματοοικονομικών συναλλαγών. Οι μηχανές μπορούν να μάθουν να αναγνωρίζουν μοτίβα στα δεδομένα και να εκχωρούν αυτόματα συναλλαγές στους αντίστοιχους λογαριασμούς στο γενικό καθολικό. Αυτό εξαλείφει την ανάγκη των λογιστών να αφιερώνουν χρόνο στην εκτέλεση αυτού του έργου χειροκίνητα, μειώνοντας έτσι τα σφάλματα και επιταχύνοντας τη διαδικασία.
- Πρόβλεψη τάσεων και πραγματοποίηση ακριβέστερων οικονομικών προβλέψεων: Αναλύοντας ιστορικά και τρέχοντα δεδομένα, οι μηχανές μπορούν να εντοπίσουν μοτίβα και να κάνουν προβλέψεις σχετικά με τη μελλοντική απόδοση μιας εταιρείας. Αυτές οι πληροφορίες μπορούν να βοηθήσουν σημαντικά στη λήψη στρατηγικών αποφάσεων και στον μακροπρόθεσμο οικονομικό σχεδιασμό (Du, 2023).

Πέρα από τα προαναφερόμενα, υπάρχουν μικρότερες εργασίες σχετικά με την λογιστική εξωτερική ανάθεση υπηρεσιών στις οποίες αξιοποιούνται ευρέως πτυχές μηχανικής μάθησης με επιδίωξη την τήρηση βιβλίων ή φορολογικών δηλώσεων. Οι σταθερές λογιστικές διεργασίες, όπως οι εισπρακτέοι λογαριασμοί και η ορθή διευθέτηση πληρωτέων λογαριασμών, η προπαρασκευή αναφορών εξόδων και η εκτίμηση κινδύνου, όλα αυτά δύναται απλά και γρήγορα να αυτοματοποιηθούν χάρη στις πολλαπλές δυνατότητες της μηχανικής μάθησης. Ενδεικτικά αναφέρεται ότι οι αλγόριθμοι μηχανικής μάθησης δύναται να συνδέσουν ένα τιμολόγιο που εισάγεται με την ανάλογη εντολή αγοράς, να προσδιορίσουν με ακρίβεια έναν λογαριασμό εξόδων αναγνώρισης και να τον εισάγουν σε ένα πλαίσιο

πληρωμών όπου ένας υπάλληλος δύναται να διερευνήσει και να καταθέσει αίτημα πληρωμής στην ουρά πληρωμών (Haddara et al., 2018).

Λόγω των πολλών θετικών που προσφέρονται, έχουν υλοποιηθεί για επενδύσεις σε πρωτοποριακές τεχνολογικές λύσεις και έχουν δημιουργήσει ορισμένες πλατφόρμες ή εργαλεία που εισάγουν αλγόριθμους μηχανικής μάθησης. Η εφαρμογή της μηχανικής μάθησης στη λογιστική μπορεί να δημιουργήσει πολλά οφέλη, μεταξύ των οποίων τα ακόλουθα (Pi, 2021):

- Μεγαλύτερη ακρίβεια: Μαθαίνοντας από ιστορικά στοιχεία και εφαρμόζοντας αυτή τη γνώση στην ανάλυση, η μηχανική μάθηση επιτρέπει πιο ακριβή και αξιόπιστα αποτελέσματα.
- Αυξημένη αποδοτικότητα: Η αυτοματοποίηση εργασιών ρουτίνας και ο εντοπισμός μοτίβων και τάσεων με αυτοματοποιημένο τρόπο επιτρέπει τη βελτιστοποίηση στις λογιστικές διαδικασίες και να μειωθεί ο χρόνος που αφιερώνεται σε μη αυτόματες εργασίες.
- Καλύτερη λήψη αποφάσεων: Η ακριβέστερη ανάλυση των λογιστικών δεδομένων που παρέχεται από τη μηχανική μάθηση παρέχει στους οικονομικούς διευθυντές σχετικές και έγκαιρες πληροφορίες για τη λήψη στρατηγικών αποφάσεων.

#### 3.4 Υπάρχουσες εφαρμογές εντοπισμού και πρόληψης της φοροδιαφυγής

Αξίζει να σημειωθεί το γεγονός ότι το 2019 το 38% των χωρών που ανήκουν στον Οργανισμό Οικονομικής Συνεργασίας και Ανάπτυξης (OECD) χρησιμοποιούσε μηχανισμούς μηχανικής μάθησης για την επιβολή της φορολογίας, ενώ ένα άλλο 34% βρισκόταν σε στάδια ανάπτυξης.

Μια εφαρμογή της μηχανικής μάθησης για τον εντοπισμό και την πρόληψη της φοροδιαφυγής είναι ότι μπορεί να ανιχνεύσει μοτίβα που θα ήταν δύσκολο να αναγνωρίσουν οι άνθρωποι. Οι αλγόριθμοι μηχανικής μάθησης μπορούν να αναλύσουν μεγάλες ποσότητες δεδομένων, συμπεριλαμβανομένων των προηγούμενων φορολογικών δηλώσεων, των τραπεζικών δηλώσεων και άλλων οικονομικών αρχείων. Με αυτόν τον τρόπο, μπορούν να εντοπίσουν ανωμαλίες και μοτίβα που μπορεί να υποδηλώνουν φοροδιαφυγή. Για παράδειγμα, εάν ένας φορολογούμενος αναφέρει με συνέπεια χαμηλότερο εισόδημα από αυτό που υποδηλώνει ο τρόπος ζωής του, αυτό μπορεί να είναι σημάδι φοροδιαφυγής.

Ωστόσο, είναι σημαντικό να σημειωθεί ότι η μηχανική μάθηση ενδέχεται να μην είναι σε θέση να εντοπίσει όλες τις περιπτώσεις φοροδιαφυγής. Η φοροδιαφυγή μπορεί να λάβει πολλές μορφές και οι φοροφυγάδες μπορεί να προσαρμοστούν σε νέες μεθόδους ανίχνευσης. Επιπλέον, οι αλγόριθμοι μηχανικής μάθησης μπορεί να κάνουν λάθη, οδηγώντας σε ψευδώς θετικά ή ψευδώς αρνητικά στοιχεία. Αυτό θα μπορούσε να έχει ως αποτέλεσμα να κατηγορηθούν νόμιμοι φορολογούμενοι για φοροδιαφυγή, η οποία θα μπορούσε να βλάψει τη φήμη τους και να οδηγήσει σε νομικές ή οικονομικές συνέπειες.

Μια ακόμα εφαρμογή της μηχανικής μάθησης σε αυτόν τον τομέα είναι η βελτίωση της συλλογής φόρων. Με την αυτοματοποίηση της διαδικασίας εντοπισμού φοροδιαφυγής, η μηχανική μάθηση μπορεί να μειώσει την ανάγκη για μη αυτόματους ελέγχους, που μπορεί να είναι χρονοβόροι και δαπανηροί. Αυτό μπορεί να απελευθερώσει πόρους για άλλες εργασίες είσπραξης φόρων, όπως η ενημέρωση και η εκπαίδευση. Επιπλέον, η μηχανική μάθηση μπορεί να βοηθήσει στον εντοπισμό της φοροδιαφυγής πιο γρήγορα και με ακρίβεια, γεγονός που μπορεί να αυξήσει τα συνολικά έσοδα που εισπράττει η κυβέρνηση.

Ωστόσο, η αυξημένη απόδοση μπορεί να έχει κόστος. Οι αλγόριθμοι μηχανικής μάθησης ενδέχεται να επισημάνουν τις νόμιμες φορολογικές δηλώσεις ως ύποπτες, οδηγώντας στην ανάγκη για μη αυτόματο έλεγχο. Αυτό θα μπορούσε να οδηγήσει σε καθυστερήσεις και πρόσθετο κόστος για τους φορολογούμενους και την κυβέρνηση. Ακόμα, μπορεί να μην λαμβάνει υπόψη ελαφρυντικές περιστάσεις, όπως ιατρικά έξοδα ή απώλεια εργασίας, κάτι που μπορεί να οδηγήσει σε προσωρινές μειώσεις του εισοδήματος. Οι ψευδείς κατηγορίες για φοροδιαφυγή θα μπορούσαν να υπονομεύσουν περαιτέρω την εμπιστοσύνη του κοινού στο φορολογικό σύστημα.

Μια τρίτη εφαρμογή είναι ότι μπορεί να μειώσει τη διαφθορά στην είσπραξη φόρων. Με τη μείωση της πιθανότητας ανθρώπινου λάθους ή μεροληψίας, η μηχανική μάθηση μπορεί να βοηθήσει να διασφαλιστεί ότι οι φόροι εισπράττονται δίκαια. Επιπλέον, η μηχανική μάθηση μπορεί να βοηθήσει στον εντοπισμό διεφθαρμένων αξιωματούχων που ενδέχεται να εμπλέκονται σε φοροδιαφυγή ή άλλες παράνομες δραστηριότητες. Αυτό μπορεί να βελτιώσει τη διαφάνεια στη συλλογή φόρων και να αυξήσει την εμπιστοσύνη του κοινού στην κυβέρνηση.

Ωστόσο, η μηχανική μάθηση μπορεί να μην αντιμετωπίσει τις βαθύτερες αιτίες της φοροδιαφυγής. Η φοροδιαφυγή μπορεί να οφείλεται σε κοινωνικοοικονομικούς παράγοντες, όπως υψηλά επίπεδα ανισότητας ή έλλειψη πρόσβασης σε χρηματοοικονομικές υπηρεσίες. Η



διαφθορά μπορεί να είναι βαθιά ενσωματωμένη στο σύστημα είσπραξης φόρων, κάτι που θα μπορούσε να απαιτεί συστημικές αλλαγές και όχι τεχνολογικές λύσεις. Η μηχανική μάθηση από μόνη της μπορεί να μην είναι αρκετή για την αντιμετώπιση αυτών των υποκείμενων ζητημάτων.

Η μηχανική μάθηση μπορεί να βοηθήσει στην αυτοματοποίηση και βελτιστοποίηση διαδικασιών όπως ο υπολογισμός φόρου, η επεξεργασία φορολογικών δηλώσεων και η διαχείριση φορολογικών δεδομένων και φορολογικών διαδικασιών: Αυτή η νέα τεχνολογική εποχή καθιστά δυνατή την αυτοματοποίηση διαδικασιών που απαιτούσαν ώρες, απελευθερώνοντας έτσι πολύτιμο χρόνο για τους επαγγελματίες λογιστές να επικεντρωθούν σε πιο στρατηγικά καθήκοντα και καθήκοντα υψηλότερης προστιθέμενης αξίας και για την αυτόματη υπολογιστική φορολογική δεδομένων και τον προσδιορισμό των φορολογικών υποχρεώσεων και των φορολογικών στοιχείων.

## Κεφάλαιο 4<sup>ο</sup>: Μελέτες περίπτωσης μηχανικής μάθησης στη φοροδιαφυγή

Η χρήση της μηχανικής μάθησης στον τομέα της φοροδιαφυγής παρέχει βελτιωμένη αποτελεσματικότητα στον εντοπισμό και στην αντιμετώπιση περιπτώσεων φοροδιαφυγής, ενισχύοντας μια πιο διαφανή και δίκαιη προσέγγιση στη φορολογική επιβολή.

Η μηχανική μάθηση αναδεικνύεται ως το βασικό μέσο τόσο για την πρόληψη όσο και για την αντιμετώπιση της φοροδιαφυγής, διαδραματίζοντας κρίσιμο ρόλο στη διατήρηση της οικονομικής ακεραιότητας και στην προώθηση της διαφάνειας στα φορολογικά συστήματα (Silveira-Santos *et al.*, 2022). Μέσω της χρήσης αλγορίθμων μηχανικής μάθησης, οι φορολογικές αρχές έχουν τη δυνατότητα να εντοπίσουν περίπλοκες και εκτεταμένες περιπτώσεις φοροδιαφυγής, οι οποίες θα μπορούσαν να παραμείνουν απαρατήρητες με τη χρήση παραδοσιακών μεθόδων ελέγχου (Dineva and Atanasova, 2020).

Η φοροδιαφυγή είναι ένα σοβαρό ζήτημα σε πολλές χώρες, συμπεριλαμβανομένης της Ελλάδας. Με τη χρήση της μηχανικής μάθησης, οι φορολογικές αρχές μπορούν να προλάβουν πιθανές παραβιάσεις και να προβούν σε εντατικούς ελέγχους σε συγκεκριμένους φορολογούμενους ή εταιρείες που παρουσιάζουν ύποπτη συμπεριφορά. Αυτό οδηγεί σε μια πιο αποτελεσματική και δικαιότερη φορολογική επιβολή, ενισχύοντας την οικονομική διαφάνεια και διατηρώντας την οικονομική ισορροπία των κρατών (Dineva and Atanasova, 2020).

### 4.1 Ελλάδα

Στην Ελλάδα, η μηχανική μάθηση έχει αρχίσει να χρησιμοποιείται σε διάφορους τομείς, συμπεριλαμβανομένων και ορισμένων που σχετίζονται με τη φορολογία. Οι αλγόριθμοι μηχανικής μάθησης έχουν αποδειχθεί χρήσιμοι στην ανάλυση οικονομικών δεδομένων για τον εντοπισμό προτύπων φοροδιαφυγής. Η Ανεξάρτητη Αρχή Δημοσίων Εσόδων (ΑΑΔΕ) έχει αναπτύξει πολλαπλές τακτικές για τον εντοπισμό και την πρόληψη της φοροδιαφυγής.

Η χρήση αλγορίθμων μηχανικής μάθησης για τον εντοπισμό προτύπων φοροδιαφυγής δεν είναι απαλλαγμένη από σφάλματα και προκαταλήψεις. Ως εκ τούτου, απαιτείται προσεκτική εφαρμογή και πολλαπλοί έλεγχοι επαλήθευσης για την επιτυχή χρήση αυτών των αλγορίθμων στα συστήματα φορολογικού ελέγχου, μαζί με την υιοθέτηση διασφαλίσεων για τη διασφάλιση της σωστής αλγοριθμικής διακυβέρνησης στους φορολογικούς ελέγχους. Οι

αλγόριθμοι μηχανικής μάθησης μπορούν να κατηγοριοποιήσουν τους φορολογούμενους σύμφωνα με τη συμπεριφορά τους, όπως ο στρατηγικός κακοπληρωτής ή η πιθανότητα φοροδιαφυγής. Τα αναλυόμενα δεδομένα μπορούν να οδηγήσουν σε στοχευμένους φορολογικούς ελέγχους και η τεχνητή νοημοσύνη μπορεί εύκολα να εντοπίσει αποκλίσεις μεταξύ των εσόδων και των δαπανών των νοικοκυριών και των επιχειρήσεων (Χατζηνικολάου, 2023).

Ακόμα, οι αλγόριθμοι μηχανικής μάθησης μπορούν να χρησιμοποιηθούν για την ανάλυση οικονομικών δεδομένων σε πραγματικό χρόνο για τον εντοπισμό ύποπτων περιπτώσεων φοροδιαφυγής και λαθρεμπορίου, όπως αποδεικνύεται από τη χρήση ηλεκτρονικής παρακολούθησης από την Ανεξάρτητη Αρχή Δημοσίων Εσόδων με τεχνητή νοημοσύνη για την παρακολούθηση των φορολογουμένων, με σχέδια για τη χρήση της περαιτέρω για τον περιορισμό της φοροδιαφυγής (Χατζηνικολάου, 2023).

Ακόμα, η ΑΑΔΕ επιδιώκει την ηλεκτρονική παρακολούθηση των φορολογουμένων για τον εντοπισμό ενδεχόμενης φοροδιαφυγής. Η στενή παρακολούθηση των φορολογουμένων περιλαμβάνει τη συλλογή δεδομένων σχετικά με το ιστορικό φορολογικής συμμόρφωσης, την τρέχουσα κατάσταση συμμόρφωσης, την οικονομική κατάσταση, την πιστοληπτική ικανότητα και τη φερεγγυότητα. Σκοπός της παρακολούθησης είναι η πρόληψη συμπεριφορών που στοχεύουν στη φοροδιαφυγή, το λαθρεμπόριο, την απάτη, τα παράνομα κέρδη και την καθυστερημένη πληρωμή φόρων. Σε περίπτωση που η πρόληψη και η αποτροπή καταστούν αδύνατες, η ΑΑΔΕ διενεργεί ελέγχους όταν οι φορολογούμενοι εκδηλώνουν παραβατικές συμπεριφορές. Η εκτίμηση κινδύνου είναι μια άλλη μέθοδος που χρησιμοποιείται για την αξιολόγηση του κινδύνου μη πληρωμής των φορολογουμένων. Το έργο περιλαμβάνει την έγκαιρη ανίχνευση περιστατικών φοροδιαφυγής σε πραγματικό χρόνο (Παλαιτσάκης, 2023).

Η ΑΑΔΕ χρησιμοποιεί πρόβλεψη εισοδήματος και οφειλών με βάση τα ιστορικά δεδομένα της συμπεριφοράς των φορολογουμένων. Επιπλέον, η ανάλυση δεδομένων των δεδομένων της ΑΑΔΕ σε συνδυασμό με εξωτερικές πηγές επιτρέπει την ενημέρωση της ανάλυσης κινδύνου, τη στόχευση και την ιεράρχηση των ελέγχων. Ανακαλύπτονται επίσης δυναμικές σχέσεις μεταξύ των φορολογουμένων για τον εντοπισμό της φοροδιαφυγής.

Τέλος, ένα ακόμα παράδειγμα στην Ελλάδα είναι το σύστημα «Οικονομική Δήλωση Φόρου Εισοδήματος» που εκμεταλλεύεται η ΑΑΔΕ. Το σύστημα αυτό χρησιμοποιεί τη μηχανική μάθηση για την επεξεργασία και ανάλυση των φορολογικών δηλώσεων των φορολογουμένων. Συγκεκριμένα, το σύστημα επιτρέπει στους φορολογούμενους να υποβάλλουν τις φορολογικές

δηλώσεις ηλεκτρονικά μέσω της πλατφόρμας της ΑΑΔΕ. Με τη χρήση της τεχνολογίας μηχανικής μάθησης, το σύστημα μπορεί να αναλύσει τα δεδομένα των δηλώσεων και να ανιχνεύσει πιθανές ανωμαλίες ή αντιφάσεις στις φορολογικές δηλώσεις. Αυτό μπορεί να περιλαμβάνει την αυτόματη ανίχνευση πιθανών περιπτώσεων φοροαπάτης ή τη μη συμμόρφωση με τη φορολογική νομοθεσία.

Με την εφαρμογή της μηχανικής μάθησης σε αυτό το πλαίσιο, η ΑΑΔΕ μπορεί να βελτιωθεί η αποτελεσματικότητα και η ακρίβεια της διαχείρισης των φορολογικών δηλώσεων, εξασφαλίζοντας την εφαρμογή της φορολογικής νομοθεσίας και την αποτροπή της φοροαπάτης.

Η εφαρμογή της μηχανικής μάθησης για την πρόληψη της φοροδιαφυγής στην Ελλάδα παρουσιάζει πιθανές προκλήσεις και περιορισμούς. Μια πιθανή πρόκληση είναι ότι οι μέθοδοι που χρησιμοποιούνται για την πρόληψη της φοροδιαφυγής θα πρέπει να παραμείνουν ανθρωποκεντρικές, με την τεχνολογία να λειτουργεί για να αυξάνει, αντί να αντικαθιστά, την ανθρώπινη εργασία. Τα μοντέλα μηχανικής μάθησης έχουν δείξει ότι είναι επιρρεπή σε εσφαλμένες εκτιμήσεις, ιδιαίτερα στο συνεχώς εξελισσόμενο φορολογικό νομικό περιβάλλον. Ως εκ τούτου, τα ευρήματα που παράγονται από συστήματα μηχανικής μάθησης δεν θα πρέπει να αντικαθιστούν το έργο των φορολογικών ελεγκτών, αλλά θα πρέπει να λειτουργούν συμπληρωματικά σε αυτό (Κρητικός, 2024).

Είναι επίσης σημαντικό να σημειωθεί ότι η εγκυρότητα των συμπερασμάτων που προκύπτουν από την εφαρμογή συστημάτων μηχανικής μάθησης για την πρόληψη της φοροδιαφυγής στην Ελλάδα πρέπει να ελέγχεται συνεχώς (Κρητικός, 2024). Αυτή η επαλήθευση μπορεί να επιτευχθεί μέσω συνεχούς παρακολούθησης, δοκιμής και ελέγχου των αλγορίθμων μηχανικής εκμάθησης για να διασφαλιστεί ότι παράγουν ακριβή αποτελέσματα. Τέλος, είναι σημαντικό να διασφαλιστεί ότι τα μοντέλα μηχανικής μάθησης που χρησιμοποιούνται για τον εντοπισμό φοροδιαφυγής στην Ελλάδα είναι διαφανή και εξηγήσιμα, έτσι ώστε οι ελεγκτές να μπορούν να κατανοήσουν πώς λαμβάνονται οι αποφάσεις και να εντοπίσουν τυχόν προκαταλήψεις ή σφάλματα.

Έτσι, στην Ελλάδα, η μηχανική μάθηση έχει αρχίσει να χρησιμοποιείται σε διάφορους τομείς, συμπεριλαμβανομένων και ορισμένων που σχετίζονται με τη φορολογία. Η κατηγοριοποίηση και το προφίλ των φορολογουμένων χρησιμοποιούνται για την εκτίμηση του κινδύνου και τον εντοπισμό μη εμφανών ομοιοτήτων για τη σύλληψη φοροφυγάδων. Αυτές οι πολλαπλές

μέθοδοι που χρησιμοποιεί η ΑΑΔΕ ήταν χρήσιμες για τον εντοπισμό και την πρόληψη της φοροδιαφυγής.

#### 4.2 Μεξικό

Οι Martin Zumaya, Rita Guerrero et.al στο άρθρο «Identifying tax evasion in Mexico with tools from network science and machine learning» που δημοσιεύθηκε το 2021 αναφέρονται στο πώς το Μεξικό βρίσκει την φοροδιαφυγή μέσα από τεχνικές μηχανικής μάθησης. Πιο αναλυτικά, το Μεξικό διατηρεί ηλεκτρονικά αρχεία όλων των φορολογητέων συναλλαγών από το 2014. Τα ανώνυμα δεδομένα που συλλέγονται από την ομοσπονδιακή κυβέρνηση του Μεξικού περιλαμβάνουν περισσότερους από 80 εκατομμύρια συνεισφέροντες (ιδιώτες και εταιρείες) και σχεδόν 7 δισεκατομμύρια μηνιαίες συγκεντρώσεις τιμολογίων μεταξύ των συνεισφερόντων μεταξύ Ιανουαρίου 2015 και Δεκεμβρίου 2018. Τα δεδομένα περιλαμβάνουν μια λίστα με σχεδόν δέκα χιλιάδες συνεισφέροντες που έχουν ήδη αναγνωριστεί ως φοροφυγάδες, λόγω των δραστηριοτήτων τους που κατασκευάζουν τιμολόγια για ανύπαρκτα προϊόντα ή υπηρεσίες, ώστε οι παραλήπτες να μπορούν να φοροδιαφύγουν. Αξιοποιώντας αυτό το εκτεταμένο σύνολο δεδομένων, χτίζουμε μηνιαία και ετήσια χρονικά δίκτυα όπου οι κόμβοι είναι οι συνεισφέροντες και οι κατευθυνόμενοι σύνδεσμοι είναι τιμολόγια που παράγονται σε ένα δεδομένο χρονικό διάστημα. Εξερευνώντας τις ιδιότητες των γειτονιών του δικτύου γύρω από φοροφυγάδες, δεικνύεται ότι τα μοτίβα αλληλεπίδρασής τους διαφέρουν από αυτά της πλειοψηφίας των συντελεστών. Ειδικότερα, οι βρόχοι τιμολόγησης μεταξύ φοροφυγάδων και πελατών τους υπερεκπροσωπούνται.

Με αυτήν τη γνώση, για να ταξινομηθούν οι συνεισφέροντες ως ύποπτοι φοροδιαφυγής χρησιμοποιούνται δύο μέθοδοι μηχανικής μάθησης: νευρωνικά δίκτυα και τυχαία δάση. Κάθε μέθοδος εκπαιδεύεται με ένα μέρος της λίστας φοροφυγάδων και τη δοκιμάζει με τις υπόλοιπες, λαμβάνοντας ακρίβεια μεγαλύτερη από 0,9 και με τις δύο μεθόδους. Χρησιμοποιώντας το πλήρες σύνολο δεδομένων των συνεισφερόντων, κάθε μέθοδος ταξινομεί περισσότερους από 100 χιλιάδες ύποπτους για φοροδιαφυγή, με περισσότερους από 40 χιλιάδες ύποπτους να ταξινομούνται και με τις δύο μεθόδους. Ο αριθμός των υπόπτων μειώνεται περαιτέρω εστιάζοντας σε όσους έχουν μικρή απόσταση δικτύου από γνωστούς φοροφυγάδες. Έτσι, εξάγεται μια λίστα με εξαιρετικά ύποπτους συνεισφέροντες ταξινομημένα με βάση το ποσό της φοροδιαφυγής, πολύτιμες πληροφορίες για τις αρχές για περαιτέρω διερεύνηση της παράνομης φορολογικής δραστηριότητας στο Μεξικό. Με την εφαρμογή

αυτών των μεθόδων υπολογίζεται η φοροδιαφυγή, που δεν είχε εντοπιστεί στο παρελθόν της τάξης των 10 δισεκατομμυρίων δολαρίων ΗΠΑ ετησίως από περίπου 10.000 συνεισφέροντες.

Η φοροδιαφυγή είναι ένα σοβαρό πρόβλημα που επηρεάζει τις κυβερνήσεις παγκοσμίως. Υπολογίζεται ότι η φοροδιαφυγή κοστίζει στην κυβέρνηση των ΗΠΑ περίπου 500 δισεκατομμύρια δολάρια ετησίως. Οι παραδοσιακές μέθοδοι ανίχνευσης και πρόληψης της φοροδιαφυγής ήταν αναποτελεσματικές, οδηγώντας στην εξερεύνηση εναλλακτικών προσεγγίσεων όπως η μηχανική μάθηση. Η μηχανική μάθηση είναι μια μορφή τεχνητής νοημοσύνης που επιτρέπει στα συστήματα υπολογιστών να μαθαίνουν και να βελτιώνονται από την εμπειρία χωρίς να είναι ρητά προγραμματισμένα. Αυτό το δοκίμιο θα διερευνήσει το ρόλο της μηχανικής μάθησης στον εντοπισμό και την πρόληψη της φοροδιαφυγής, τα οφέλη, τους περιορισμούς και τους ηθικούς λόγους.

#### 4.3 Ισπανία

Η μηχανική μάθηση μπορεί επίσης να χρησιμοποιηθεί για την πρόληψη της φοροδιαφυγής βελτιώνοντας τη φορολογική συμμόρφωση. Για παράδειγμα, οι αλγόριθμοι μηχανικής μάθησης μπορούν να χρησιμοποιηθούν για την ανάλυση δεδομένων φορολογουμένων και τον εντοπισμό περιοχών μη συμμόρφωσης. Αυτές οι πληροφορίες μπορούν στη συνέχεια να χρησιμοποιηθούν για την ανάπτυξη στοχευμένων προγραμμάτων συμμόρφωσης που αντιμετωπίζουν συγκεκριμένα ζητήματα. Επιπλέον, οι αλγόριθμοι μηχανικής μάθησης μπορούν να χρησιμοποιηθούν για τον εντοπισμό φορολογουμένων υψηλού κινδύνου που είναι πιο πιθανό να εμπλακούν σε συμπεριφορές που δεν συμμορφώνονται, επιτρέποντας στις φορολογικές αρχές να εστιάζουν τους πόρους τους σε αυτά τα άτομα. Η χρήση της μηχανικής μάθησης στη φορολογική συμμόρφωση έχει ήδη δείξει πολλά υποσχόμενα αποτελέσματα. Για παράδειγμα, η ισπανική φορολογική αρχή, Agencia Tributaria, χρησιμοποιεί αλγόριθμους μηχανικής εκμάθησης για να αναλύσει τα δεδομένα των φορολογουμένων και να εντοπίσει περιοχές μη συμμόρφωσης. Αυτή η προσέγγιση οδήγησε σε αυξημένα φορολογικά έσοδα και βελτιωμένα ποσοστά συμμόρφωσης. Ωστόσο, υπάρχουν πιθανοί περιορισμοί και προκλήσεις στη χρήση της μηχανικής μάθησης για την πρόληψη της φοροδιαφυγής. Μία από τις προκλήσεις είναι η ανάγκη για ακριβή δεδομένα για την εκπαίδευση των αλγορίθμων. Επιπλέον, υπάρχει κίνδυνος υπερβολικής εξάρτησης από αλγόριθμους μηχανικής μάθησης, γεγονός που μπορεί να οδηγήσει στην παραμέληση άλλων σημαντικών παραγόντων που συμβάλλουν στη φορολογική συμμόρφωση.

#### 4.4 Ινδία

Στο άρθρο των Ankit Rathi, Dr. Saurabh Sharma, Dr. Gaurav Lodha, Dr. Manoj Srivastava «A Study on Application of Artificial Intelligence and Machine Learning in Indian Taxation System» που δημοσιεύτηκε το 2021 αναφέρει ότι η φορολογία είναι η κύρια πηγή δημόσιας χρηματοδότησης σε μια αναπτυσσόμενη οικονομία όπως η Ινδία. Το ινδικό φορολογικό σύστημα εμφανίζει πολλά προβλήματα φοροδιαφυγή. Η διοίκηση της φορολογίας χρειάζεται ένα σύστημα που θα είναι λιγότερα λάθη και θα είναι γρήγορο στη λήψη αποφάσεων. Πάσχει από έλλειψη ανθρώπινου δυναμικού για την εκτέλεση κουραστικών εργασιών, όπως η εισαγωγή δεδομένων, ο έλεγχος των δηλώσεων. Η κυβέρνηση της Ινδίας ανακοίνωσε ότι θα χρησιμοποιήσει απρόσωπο σύστημα φορολογικής εκτίμησης που εξουσιοδοτείται από τεχνητή νοημοσύνη και μηχανική μάθηση. Οι παράγοντες που παρατηρούνται είναι η φορολογική γνώση, η φορολογική εκπαίδευση, η νομική κύρωση, η πολυπλοκότητα του φορολογικού συστήματος, η σχέση με τη φορολογική αρχή, η αντιληπτή δικαιοσύνη των φορολογικό σύστημα, ηθική και στάση απέναντι στη φορολογική συμμόρφωση.

Η μελέτη έδειξε ότι παρόλο που οι φορολογούμενοι παρατηρούν ότι είναι δική τους ευθύνη να πληρώνουν φόρους, η Τεχνητή Νοημοσύνη θα έχει μικρότερο αντίκτυπο στην εμπειρία πλήρωσης φόρων και μπορούμε να πούμε ότι υπάρχει λιγότερη ευαισθητοποίηση σχετικά με την ίδια την έννοια. Η κυβέρνηση κάνει πολλές δαπάνες για να ενσωματωθεί το φορολογικό σύστημα με την τεχνητή νοημοσύνη, αλλά εάν οι φορολογούμενοι δεν γνωρίζουν το σύστημα, τότε θα είναι δύσκολο να επιτύχει αυτή η μεταρρύθμιση. Η κυβέρνηση απαιτεί ένα σύστημα που βασίζεται στην τεχνητή νοημοσύνη και οι φορολογούμενοι είναι έτοιμοι για την επερχόμενη τεχνολογική αλλαγή στο σύστημα, αλλά δεν θέλουν να τερματίσουν την ανθρώπινη παρέμβαση. Οι φορολογούμενοι πιστεύουν ακράδαντα ότι η χρήση ενός φορολογικού συστήματος που βασίζεται στην τεχνητή νοημοσύνη θα είναι ασφαλής, ασφαλές και άνετο στη χρήση. Εάν οι φορολογούμενοι εμπιστεύονται ολόκληρο το φορολογικό σύστημα, τότε μόνο αυτοί θα παράσχουν τα σωστά δεδομένα και γι' αυτό θα μπορέσουμε να αποκτήσουμε ένα πολύ ισχυρό σύστημα φορολογικής διοίκησης, αλλά γι' αυτό η κυβέρνηση πρέπει να ενημερώσει τους ανθρώπους για το σύστημα και να κάνει το σύστημα διαφανές. ότι θα οικοδομήσει την εμπιστοσύνη των φορολογουμένων και θα το δεχτούν ως βοηθητικό εργαλείο για ένα δίκαιο και διαφανές ινδικό φορολογικό σύστημα.

#### 4.5 Τουρκία

Οι Hamal & Senvar στο άρθρο «Comparing performances and effectiveness of machine learning classifiers in detecting financial accounting fraud for Turkish SMEs» που δημοσιεύτηκε το 2021, μελέτησαν την αποτελεσματικότητα των τεχνικών μηχανικής μάθησης για την εύρεση επιχειρήσεων που προσπάθησαν να εξαπατήσουν το κράτος μέσω των βασικών τους οικονομικών καταστάσεων (Hamal and Senvar, 2021). Σε αυτό το πλαίσιο οι συγγραφείς υλοποίησαν συγκεκριμένα πειράματα αξιοποιώντας χαρακτηριστικούς αλγορίθμους μάθησης όπου παρατήρησαν ότι η ενδεδειγμένη μεθοδολογία παραλλαγής στοίβαξης υλοποιήθηκε πιο σωστά συγκριτικά με άλλες μεθόδους. Πιο συγκεκριμένα, οι τουρκικές μικρομεσαίες επιχειρήσεις είναι εκτεθειμένες σε κινδύνους απάτης και οι πιστώτριες τράπεζες αντιμετωπίζουν μεγάλες προκλήσεις για την αντιμετώπιση της απάτης στη χρηματοοικονομική λογιστική. Διερευνούν την αποτελεσματικότητα των ταξινομητών μηχανικής μάθησης στην ανίχνευση της χρηματοοικονομικής λογιστικής απάτης αξιολογώντας τις οικονομικές καταστάσεις 341 τουρκικών ΜΜΕ από το 2013 έως το 2017. Τα δεδομένα πάρθηκαν από κάποια τράπεζα της Τουρκίας. Αρχικά, έγινε η προεπεξεργασία των δεδομένων, όπου πραγματοποιήθηκε ο υπολογισμός των χρηματοοικονομικών δεικτών, οι μέθοδοι επιλογής χαρακτηριστικών για τον καθορισμό των χρηματοοικονομικών δεικτών με τη μεγαλύτερη επίδραση στις δόλιες οικονομικές καταστάσεις και δύο μέθοδοι δειγματοληψίας της Synthetic Minority Oversampling Technique (SMOTE) ως υπερδειγματοληψία. Μετά έγινε η αξιολόγηση των επιδόσεων και η σύγκριση των ταξινομητών όπου εκτελούνται επτά διαφορετικοί ταξινομητές (μηχανή διανυσμάτων υποστήριξης, Naive Bayes, τεχνητό νευρωνικό δίκτυο, K-κοντινότερος γείτονας, τυχαίο δάσος, λογιστική παλινδρόμηση και bagging) και συγκρίνονται με τη χρήση μετρικών επιδόσεων. Οι ταξινομητές συγκρίνονται επίσης χωρίς τη χρήση τεχνικών επιλογής χαρακτηριστικών ή/και δειγματοληψίας. Τα αποτελέσματα αποκαλύπτουν ότι το μοντέλο τυχαίου δάσους χωρίς επιλογή χαρακτηριστικών-υπερδειγματοληψία υπερτερεί έναντι όλων των άλλων μοντέλων.

#### 4.6 Πολωνία

Το σύστημα Teleinformatyczny Izby Rozliczeniowej (STIR) της Πολωνίας είναι ένα τεχνικό σύστημα που στοχεύει στον αυτόματο εντοπισμό ύποπτων συναλλαγών σε άμεσο χρόνο.

Το Πολωνικό σύστημα συλλέγει στοιχεία από τους τραπεζικούς λογαριασμούς που έχουν οι εταιρίες, συμπεριλαμβανομένων των ημερήσιων καταστάσεων συναλλαγών, των στοιχείων αναγνώρισης των αποστολέων και των παραληπτών των συναλλαγών και των αριθμών



λογαριασμού τους. Συλλέγει επίσης, την ημερομηνία, το ποσό, το νόμισμα, τον τίτλο και την περιγραφή των συναλλαγών. Σημειώνεται ότι από τον Ιούλιο του 2019, οι τράπεζες πρέπει επίσης να παρέχουν στο σύστημα STIR τις διευθύνσεις IP από τις οποίες οι κάτοχοι λογαριασμού συνδέονται στο τραπεζικό σύστημα και διενεργούν τις συναλλαγές τους.

Τα στοιχεία που έδωσε στη δημοσιότητα η ΚΑΣ, η Εθνική Διοίκηση Εσόδων, έδειξαν ότι το STIR συνέλεξε στοιχεία για περισσότερες από 11 εκατομμύρια συναλλαγές το 2019, που αφορούσαν σχεδόν 4 εκατομμύρια οντότητες.

## Συμπεράσματα και Προτάσεις

Η σύνδεση της μηχανικής μάθησης με τη φοροδιαφυγή σημαίνει ότι οι υπολογιστές μπορούν να μάθουν από φορολογικά δεδομένα και να βοηθήσουν στον εντοπισμό παραβάσεων φορολογικών κανόνων. Αυτό μπορεί να βοηθήσει να βρεθούν φορολογούμενοι που προσπαθούν να αποφύγουν την πληρωμή των φόρων τους. Ο εντοπισμός πρότυπων ή συμπεριφορών που υποδηλώνουν πιθανή φοροδιαφυγή, μπορεί να γίνει μέσω διάφορων τεχνικών μηχανικής μάθησης, όπως η επιβλεπόμενη μάθηση και η μη επιβλεπόμενη μάθηση. Αυτό τους επιτρέπει να εκτελούν πολύπλοκες εργασίες με αυξημένη αποτελεσματικότητα και ακρίβεια. Έτσι μπορούν να αναλύουν τεράστιες ποσότητες φορολογικών δεδομένων και να εντοπίζουν μοτίβα που υποδηλώνουν πιθανές περιπτώσεις φοροδιαφυγής. Με την αξιοποίηση αυτών των μοντέλων οι φορολογικές αρχές μπορούν να προβούν σε εξειδικευμένους ελέγχους και έρευνες για την αντιμετώπιση πιθανών περιστατικών της φοροδιαφυγής. Οι αλγόριθμοι μηχανικής μάθησης χρησιμοποιούνται ευρέως για τον εντοπισμό προτύπων και ανωμαλιών σε δεδομένα φορολογικών δηλώσεων, επιτρέποντας την ανίχνευση φοροφυγάδων και την πρόβλεψη μη συμμόρφωσης.

Μέσω αυτών των προηγμένων αλγορίθμων και την ανάλυση δεδομένων, η μηχανική μάθηση έχει αρχίσει να διαδραματίζει σημαντικό ρόλο στην αντιμετώπιση της φοροδιαφυγής, προσφέροντας ανίχνευση παραβατών, ανιχνεύοντας μοτίβα φοροδιαφυγής και ανωμαλιών στα φορολογικά δεδομένα, πρόβλεψη φορολογικών εσόδων, επιτρέποντας στις φορολογικές αρχές να προετοιμάζονται και να προγραμματίζουν κατάλληλα τους πόρους τους, αυτοματοποίηση ελέγχων και εξατομίκευση της επιβολής προσαρμόζοντας την επιβολή των φόρων σύμφωνα με τα ειδικά χαρακτηριστικά και τις συνήθειες κάθε φορολογίας.

Συνολικά, η μηχανική μάθηση μπορεί να παρέχει σημαντικά οφέλη στην επιβολή και τη συλλογή φόρων, αλλά απαιτεί προσεκτική χρήση και εποπτεία για να αποφευχθούν πιθανά προβλήματα και να αντιμετωπιστούν οι περιορισμοί της, όπως η εξάρτηση της ακρίβειας από την ποιότητα των δεδομένων, η ηθική ανησυχία για τη διαφάνεια και την πιθανή μεροληπτικότητα των αλγορίθμων, ο κίνδυνος ψευδός θετικών ή αρνητικών αποτελεσμάτων. Τέλος, η μηχανική μάθηση μπορεί να μην αντιμετωπίσει τις αιτίες της φοροδιαφυγής.

Οι τεχνικές μηχανικής μάθησης, είτε με επίβλεψη είτε χωρίς, έχουν φέρει σημαντική βελτίωση στην αποτελεσματικότητα και ακρίβεια των φορολογικών ελέγχων. Μια μελέτη από την Ιταλία

αναφέρει αύξηση της ανιχνευμένης φοροδιαφυγής και μείωση του ποσού της, χάρη στην αντικατάσταση των ελέγχων από εκπαιδευμένους αλγορίθμους.

Σε ό,τι αφορά την Ελλάδα, η μηχανική μάθηση έχει αρχίσει να εφαρμόζεται επιτυχώς σε διάφορους τομείς της φορολογίας, παρέχοντας σημαντικά οφέλη στον εντοπισμό και την πρόληψη της φοροδιαφυγής. Διάφορες πτυχές της εφαρμογής τεχνικών μηχανικής μάθησης είναι ο εντοπισμός προτύπων φοροδιαφυγής, η ανάλυση οικονομικών δεδομένων σε πραγματικό χρόνο, η παρακολούθηση των φορολογουμένων και η επεξεργασία φορολογικών δηλώσεων.

Εν κατακλείδι, η χρήση μηχανικής μάθησης και φορολογικών δεδομένων με τη σωστή επιλογή αλγόριθμου μηχανικής μάθησης έχει αποτελεσματικότερες και πιο δίκαιες διαδικασίες επιβολής φορών, βοηθώντας τις φορολογικές αρχές να διατηρήσουν τη φορολογική δικαιοσύνη και την οικονομική σταθερότητα, αλλά και ελέγχου και πρόβλεψης της φορολογικής απάτης. Παρά τα οφέλη, υπάρχουν προκλήσεις και περιορισμοί που πρέπει να ληφθούν υπόψη, όπως η ανάγκη για ανθρωποκεντρική προσέγγιση και η διασφάλιση της διαφάνειας και εξηγησιμότητας των μοντέλων μηχανικής μάθησης. Με τη σωστή επιλογή αλγόριθμου μηχανικής μάθησης ο εντοπισμός της φοροδιαφυγής θα είναι πιο αποτελεσματικός.

Η συχνότητα και οι προκλήσεις που εγκυμονούν στο φαινόμενο της φοροδιαφυγής, καθώς και η ανάγκη για άμεσες και αποτελεσματικές λύσεις ενισχύουν τη χρήση της μηχανικής μάθησης στον τομέα αυτό. Αυτή η σύνδεση ανοίγει νέους τρόπους εντοπισμού και πρόληψης της φοροδιαφυγής, αλλά απαιτεί προσοχή στην ποιότητα των δεδομένων και τις ηθικές επιπτώσεις της χρήσης τεχνολογίας.

## Βιβλιογραφία

- Adams, C. (2006) *For Good and Evil: The Impact of Taxes on the Course of Civilization*. Madison Books (G - Reference, Information and Interdisciplinary Subjects Series). Available at: <https://books.google.gr/books?id=JKKzAAAAIAAJ>.
- Aghghaleh, S.F., Mohamed, Z.M. and Rahmat, M.M. (2016) 'Detecting Financial Statement Frauds in Malaysia: Comparing the Abilities of Beneish and Dechow Models', *Asian Journal of Accounting and Governance*, 7, pp. 57–65. Available at: <https://doi.org/10.17576/AJAG-2016-07-05>.
- Ali, S. et al. (2020) '6G White Paper on Machine Learning in Wireless Communication Networks'. arXiv. Available at: <http://arxiv.org/abs/2004.13875> (Accessed: 18 March 2024).
- Alkayid, dr. K. and Ali, M. (2018) 'Applications of Big Data Analytics in Financial Auditing- A Case Study of Taiwanese Audit Firms', in.
- Alm, J. and Martinez-Vazquez (2003) 'Spatially and Persistence in U.S. Individual Income Tax Compliance'.
- Alpaydin, E. (2010) *Introduction to machine learning*. 2nd ed. Cambridge, Mass: MIT Press (Adaptive computation and machine learning).
- Alzubi, J., Nayyar, A. and Kumar, A. (2018) 'Machine Learning from Theory to Algorithms: An Overview', *Journal of Physics: Conference Series*, 1142, p. 012012. Available at: <https://doi.org/10.1088/1742-6596/1142/1/012012>.
- Appelbaum, D. et al. (2017) 'Impact of business analytics and enterprise systems on managerial accounting', *International Journal of Accounting Information Systems*, 25, pp. 29–44. Available at: <https://doi.org/10.1016/j.accinf.2017.03.003>.
- Baghdasaryan, V. et al. (2022) 'Improving Tax Audit Efficiency Using Machine Learning: The Role of Taxpayer's Network Data in Fraud Detection', *Applied Artificial Intelligence*, 36(1), p. 2012002. Available at: <https://doi.org/10.1080/08839514.2021.2012002>.
- Bao, Y. et al. (2020) 'Detecting Accounting Fraud in Publicly Traded U.S. Firms Using a Machine Learning Approach', *Journal of Accounting Research*, 58(1), pp. 199–235. Available at: <https://doi.org/10.1111/1475-679X.12292>.

- Bastanlar, Y. and Ozuysal, M. (2014) 'Introduction to Machine Learning', *Methods in molecular biology (Clifton, N.J.)*, 1107, pp. 105–128. Available at: [https://doi.org/10.1007/978-1-62703-748-8\\_7](https://doi.org/10.1007/978-1-62703-748-8_7).
- Battaglini, M. *et al.* (2022) 'Refining Public Policies with Machine Learning: The Case of Tax Auditing'.
- Bhattacharya, C. and Sinha, M. (2022) 'The Role of Artificial Intelligence in Banking for Leveraging Customer Experience', *Australasian Business, Accounting and Finance Journal*, 16(5), pp. 89–105. Available at: <https://doi.org/10.14453/aabfj.v16i5.07>.
- Bishop, C.M. (2006) *Pattern recognition and machine learning*. New York: Springer (Information science and statistics).
- Blum, A., Hopcroft, J. and Kannan, R. (2018) 'Foundations of Data Science'.
- Bolton, C. *et al.* (2018) 'The power of human–Machine collaboration: Artificial intelligence, business automation, and the smart economy', *Economics, Management, and Financial Markets*, 13(4), pp. 51–56. Available at: <https://doi.org/10.22381/EMFM13420184>.
- Brown, N.J. (1994) 'Who abolished corvée labour in Egypt and why?', *Past & Present*, (144), pp. 116–137.
- Brownlee, J. (2013) 'What is Machine Learning?', *Machine Learning Mastery*. Available at: <https://machinelearningmastery.com/what-is-machine-learning/>.
- Carbonell, J.G., Michalski, R.S. and Mitchell, T.M. (1983) *An Overview Of Machine Learning*. Available at: [https://link.springer.com/chapter/10.1007/978-3-662-12405-5\\_1](https://link.springer.com/chapter/10.1007/978-3-662-12405-5_1).
- Catak, F.O. *et al.* (2021) 'Security Concerns on Machine Learning Solutions for 6G Networks in mmWave Beam Prediction'. arXiv. Available at: <http://arxiv.org/abs/2105.03905> (Accessed: 27 October 2023).
- Cervantes, J. *et al.* (2020) 'A comprehensive survey on support vector machine classification: Applications, challenges and trends', *Neurocomputing*, 408, pp. 189–215. Available at: <https://doi.org/10.1016/j.neucom.2019.10.118>.
- Cho, S. *et al.* (2020) 'Learning from Machine Learning in Accounting and Assurance', *Journal of Emerging Technologies in Accounting*, 17(1), pp. 1–10. Available at: <https://doi.org/10.2308/jeta-10718>.

Cioffi, R. *et al.* (2019) 'Artificial Intelligence and Machine Learning Applications in Smart Production: Progress, Trends and Directions', *Sustainability*, 12(2), p. 492.

Cowell, F. (1985) 'The Economic Analysis of Tax Evasion', *Bulletin of Economic Research*, 37(3), pp. 163–193. Available at: <https://doi.org/10.1111/j.1467-8586.1985.tb00192.x>.

Cowell, F.A. (1990) 'Cheating the government: The economics of evasion', *MIT Press Books*, 1.

Damerji, H. and Salimi, A. (2021) 'Mediating effect of use perceptions on technology readiness and adoption of artificial intelligence in accounting', *Accounting Education*, 30(2), pp. 107–130. Available at: <https://doi.org/10.1080/09639284.2021.1872035>.

Dean, J. (2014) *Big Data, Data Mining, and Machine Learning*. John Wiley & Sons.

Dineva, K. and Atanasova, T. (2020) 'Systematic Look at Machine Learning Algorithms - advantages, disadvantages and practical applications', *International Multidisciplinary Scientific GeoConference: SGEM*, 20, pp. 317–324. Available at: <https://doi.org/10.5593/sgem2020/2.1/s07.041>.

Dong, H. *et al.* (2021) 'Next-Generation Data Center Network Enabled by Machine Learning: Review, Challenges, and Opportunities', *IEEE Access*, 9, pp. 136459–136475. Available at: <https://doi.org/10.1109/ACCESS.2021.3117763>.

Du, J. *et al.* (2020) 'Machine Learning for 6G Wireless Networks: Carry-Forward-Enhanced Bandwidth, Massive Access, and Ultrareliable/Low Latency', *IEEE Vehicular Technology Magazine*, 15(4), pp. 122–134. Available at: <https://doi.org/10.1109/MVT.2020.3019650>.

Du, X. (2023) 'Accounting for Machine Learning Prediction Errors in Design', *Journal of Mechanical Design*, 146(5).

El Naqa, I. and Murphy, M.J. (2015) 'What Is Machine Learning?', in I. El Naqa, R. Li, and M.J. Murphy (eds) *Machine Learning in Radiation Oncology: Theory and Applications*. Cham: Springer International Publishing, pp. 3–11. Available at: [https://doi.org/10.1007/978-3-319-18305-3\\_1](https://doi.org/10.1007/978-3-319-18305-3_1).

Elmachtoub, A.N., Liang, J.C.N. and McNellis, R. (2020) 'Decision Trees for Decision-Making under the Predict-then-Optimize Framework', *International conference on machine learning* [Preprint].

Eur-Lex (2022) *Access the authentic Official Journal of the European Union, EU law, EU case-law, consolidated texts, summaries of legislation, and much more.*

Fossaceca, J.M. and Young, S.H. (2018) 'Artificial intelligence and machine learning for future army applications', in T. Pham, M.A. Kolodny, and D.M. Wiegmann (eds) *Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR IX*. *Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR IX*, Orlando, United States: SPIE, p. 6. Available at: <https://doi.org/10.1117/12.2307753>.

Franzoni, L.A. (1999) *Tax Evasion and Tax Compliance in B. Bouckaert and G. De Geest (eds)*. USA: Edward Elgar (Encyclopedia of Law and Economics).

van den Goorbergh, R. *et al.* (2022) 'The harm of class imbalance corrections for risk prediction models: illustration and simulation using logistic regression', *Journal of the American Medical Informatics Association*, 29(9), pp. 1525–1534. Available at: <https://doi.org/10.1093/jamia/ocac093>.

Haddara, M. *et al.* (2018) 'Applications of Big Data Analytics in Financial Auditing- A Study on The Big Four', in *Americas Conference on Information Systems*. Available at: <https://api.semanticscholar.org/CorpusID:53046785>.

Hamal, S. and Senvar, O. (2021) 'Comparing performances and effectiveness of machine learning classifiers in detecting financial accounting fraud for Turkish SMEs', *International Journal of Computational Intelligence Systems*, 14(1), p. 769. Available at: <https://doi.org/10.2991/ijcis.d.210203.007>.

Helm, J.M. *et al.* (2020) 'Machine Learning and Artificial Intelligence: Definitions, Applications, and Future Directions', *Current Reviews in Musculoskeletal Medicine*, 13(1), pp. 69–76. Available at: <https://doi.org/10.1007/s12178-020-09600-8>.

Herbrich, R. and Graepel, T. (2015) 'Machine Learning. An Algorithm Perspective.'

Hornuf, L. (ed.) (2022) *Diginomics Research Perspectives: The Role of Digitalization in Business and Society*. Cham: Springer International Publishing (Advanced Studies in Diginomics and Digitalization). Available at: <https://doi.org/10.1007/978-3-031-04063-4>.

Hoydis, J. *et al.* (2020) *Towards a 6G AI-Native Air Interface*.

Jijo, B.T. and Abdulazeez, A.M. (2021) 'Classification Based on Decision Tree Algorithm for Machine Learning', *Journal of Applied Science and Technology Trends*, 2, pp. 20–28. Available at: <https://doi.org/10.38094/jastt20165>.

- Kato, N. *et al.* (2020) 'Ten challenges in advancing machine learning technologies toward 6G', 27(3), pp. 96–103.
- Kaur, J. *et al.* (2021) 'Machine Learning Techniques for 5G and Beyond', *IEEE Access*, 9, pp. 23472–23488. Available at: <https://doi.org/10.1109/ACCESS.2021.3051557>.
- Lahann, J., Scheid, M. and Fettke, P. (2019) 'Utilizing Machine Learning Techniques to Reveal VAT Compliance Violations in Accounting Data', in *2019 IEEE 21st Conference on Business Informatics (CBI)*. *2019 IEEE 21st Conference on Business Informatics (CBI)*, Moscow, Russia: IEEE, pp. 1–10. Available at: <https://doi.org/10.1109/CBI.2019.00008>.
- Langley, P. (1996) *Elements of Machine Learning*. Elsevier Science & Technology Books (Machine Learning Series). Available at: <https://books.google.gr/books?id=TNg5qVoqRtUC>.
- Lovén, L. *et al.* (2019) *EdgeAI: A Vision for Distributed, Edge-native Artificial Intelligence in Future 6G Networks*. Finland.
- Marsland, S. (2014) *Machine Learning: An Algorithmic Perspective*. 2nd edn. Chapman and Hall/CRC. Available at: <https://doi.org/10.1201/b17476>.
- Mitchell, T. *et al.* (no date) 'Machine Learning', *Annu. Rev. Comput. Sci.* [Preprint].
- Mitchell, T.M. (1997) *Machine Learning*. New York: McGraw-Hill (McGraw-Hill series in computer science).
- msn (2023) 'Η εφορία αποκτά υπερσύγχρονο σύστημα εξόρυξης δεδομένων'.
- Murorunkwere, B.F. *et al.* (2023) 'Predicting tax fraud using supervised machine learning approach', *African Journal of Science, Technology, Innovation and Development*, 15(6), pp. 731–742. Available at: <https://doi.org/10.1080/20421338.2023.2187930>.
- Nathan, J. (1994) *Who abolished corvee labour in egypt and why?* Past & Present. Available at: <https://doi.org/10.1093/past/114>.
- Nawaz, S.J. *et al.* (2019) 'Quantum Machine Learning for 6G Communication Networks: State-of-the-Art and Vision for the Future', *IEEE Access*, 7, pp. 46317–46350. Available at: <https://doi.org/10.1109/ACCESS.2019.2909490>.
- Nayak, S. and Patgiri, R. (2021) '6G Communication Technology: A Vision on Intelligent Healthcare', in, pp. 1–18. Available at: [https://doi.org/10.1007/978-981-15-9735-0\\_1](https://doi.org/10.1007/978-981-15-9735-0_1).



Pi, Y. (2021) 'Machine learning in Governments: Benefits, Challenges and Future Directions', *JeDEM - eJournal of eDemocracy and Open Government*, 13(1), pp. 203–219. Available at: <https://doi.org/10.29379/jedem.v13i1.625>.

Pisner, D. and Schnyer, D. (2020) 'Support vector machine', in *Machine Learning: Methods and Applications to Brain Disorders*. Elsevier, pp. 101–121. Available at: <https://doi.org/10.1016/B978-0-12-815739-8.00006-7>.

Poole, D.L., Mackworth, A.K. and Goebel, R. (1998) *Computational intelligence: a logical approach*. New York: Oxford University Press.

Rathore, M.M. et al. (2021) 'The Role of AI, Machine Learning, and Big Data in Digital Twinning: A Systematic Literature Review, Challenges, and Opportunities', *IEEE Access*, 9, pp. 32030–32052. Available at: <https://doi.org/10.1109/ACCESS.2021.3060863>.

Sandmo, A. (2005) 'The theory of tax evasion: A retrospective view', *National tax journal*, 58(4), pp. 643–663.

Schonlau, M. and Zou, R. (2020) 'The random forest algorithm for statistical learning', *The Stata Journal: Promoting communications on statistics and Stata*, 20(1), pp. 3–29. Available at: <https://doi.org/10.1177/1536867X20909688>.

Shanahan, J.G. (2000) 'Machine Learning', in *Soft Computing for Knowledge Discovery: Introducing Cartesian Granule Features*. Boston, MA: Springer US, pp. 143–175. Available at: [https://doi.org/10.1007/978-1-4615-4335-0\\_7](https://doi.org/10.1007/978-1-4615-4335-0_7).

Silveira-Santos, T. et al. (2022) 'Were ride-hailing fares affected by the COVID-19 pandemic? Empirical analyses in Atlanta and Boston', *Transportation*, pp. 1–32.

Slemrod, J. and Yitzhaki, S. (2002) 'Tax avoidance, evasion, and administration', in *Handbook of public economics*. Elsevier, pp. 1423–1470.

Smith, A. (1776) *An Inquiry Into the Nature and Causes of the Wealth of Nations*. London.

Sosa Hernández, V. et al. (2021) 'A Practical Tutorial for Decision Tree Induction: Evaluation Measures for Candidate Splits and Opportunities', *ACM Computing Surveys*, 54, pp. 1–38. Available at: <https://doi.org/10.1145/3429739>.

Sun, Y. et al. (2020) 'When Machine Learning Meets Privacy in 6G: A Survey', *IEEE Communications Surveys & Tutorials*, 22(4), pp. 2694–2724.

Tan, C. *et al.* (2018) 'A Survey on Deep Transfer Learning'. arXiv. Available at: <http://arxiv.org/abs/1808.01974> (Accessed: 25 March 2024).

Tang, F. *et al.* (2020) 'Future Intelligent and Secure Vehicular Network Toward 6G: Machine-Learning Approaches', *Proceedings of the IEEE*, 108(2), pp. 292–307.

Üçoğlu, D. (2020) 'Current machine learning applications in accounting and auditing', *Pressacademia*, 12(1), pp. 1–7. Available at: <https://doi.org/10.17261/Pressacademia.2020.1337>.

Visitpanya, N. and Samanchuen, T. (2023) 'A Machine Learning Approach to Identifying Suspicious Tax Evasion Behavior in Public Financial Data', in *2023 8th International Conference on Business and Industrial Research (ICBIR)*. *2023 8th International Conference on Business and Industrial Research (ICBIR)*, Bangkok, Thailand: IEEE, pp. 1152–1158. Available at: <https://doi.org/10.1109/ICBIR57571.2023.10147479>.

Wang, H. and Yeung, D.-Y. (2021) 'A Survey on Bayesian Deep Learning'. arXiv. Available at: <http://arxiv.org/abs/1604.01662> (Accessed: 25 March 2024).

Weil, S., De Silva, T.-A. and Ward, M. (2014) 'Blended learning in accounting: A New Zealand case', *Meditari Accountancy Research*, 22(2), pp. 224–244. Available at: <https://doi.org/10.1108/MEDAR-10-2013-0044>.

Yang, G. and Wang, X.-J. (2020) 'Artificial Neural Networks for Neuroscientists: A Primer', *Neuron*, 107(6), pp. 1048–1070. Available at: <https://doi.org/10.1016/j.neuron.2020.09.005>.

Zhang, Z. *et al.* (2019) '6G Wireless Networks: Vision, Requirements, Architecture, and Key Technologies', *IEEE Vehicular Technology Magazine*, 14(3), pp. 28–41. Available at: <https://doi.org/10.1109/MVT.2019.2921208>.

Ανδρεάδης, Α. (1903) 'Εισιτήριοις λόγος εις το μάθημα της δημοσιολογίας'.

Ανδρώνης, Χ. (2021) *Τεχνητή νοημοσύνη και μηχανική μάθηση στα ασύρματα δίκτυα επόμενης γενιάς*. Πανεπιστήμιο Πειραιώς.

Αργυράκης, Π. (2001) *Νευρωνικά Δίκτυα και Εφαρμογές*. ΕΛΛΗΝΙΚΟ ΑΝΟΙΚΤΟ ΠΑΝΕΠΙΣΤΗΜΙΟ. Πάτρα.

Γεωργακόπουλος, Θ. (2012) *Εισαγωγή στη δημόσια οικονομική*. Αθήνα: Μπένου.

Γεωργακόπουλος, Θ. (2016) 'Δημόσια Διοίκηση, Πολυνομία και Κακονομία στην Ελλάδα'.

Γεωργούλη, Α. (2015) *Τεχνητή Νοημοσύνη*.

Θεοχαράκης (2018) *Έρευνα για τη φύση και τις αιτίες του Πλούτου των Εθνών*. Αθήνα: Πεδίο.

Ίδρυμα Τριανταφυλλίδη (1998) *Λεξικό της κοινής νεοελληνικής*. Available at: [greek-language.gr](http://greek-language.gr).

Κούτσης, Χ. (2022) *Μηχανική μάθηση σε ιατρικά δεδομένα*. Πανεπιστήμιο Δυτικής Αττικής.

Μουστάκης, Β. (2010) *Το πρόβλημα της κατάταξης μέσα από τη χρήση μοντέλων δεσμευμένης πιθανότητας*. Πολυτεχνίο Κρήτης.

Παλαιτσάκης, Γ. (2023) *‘Τεχνητή νοημοσύνη κατά... φορολογουμένων’*, *Ναυτεμπορική*.

Παπακωνσταντίνου, Ε. and Καϊτσα, Γ. (1997) *Στατιστική*. Ευγενίδιο Ίδρυμα.

Παπαποστόλου, Σ. (2017) *Κατηγοριοποίηση με μηχανές διανυσμάτων υποστήριξης*.

Αριστοτέλειο Πανεπιστήμιο.

Πειβάνη, Λ. (2020) *Η επιστήμη της μηχανικής μάθησης*. Πανεπιστήμιο Δυτικής Αττικής.

Τεγόπουλος and Φυτράκης (1990) *Ελληνικό Λεξικό*. Δ. Αθήνα: Αρμονία.

Χατζηνικολάου, Π. (2023) *‘Έντοπισμός φοροφυγάδων με τεχνητή νοημοσύνη’*, *Καθημερινή*.