



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ
ΣΧΟΛΗ ΔΙΟΙΚΗΤΙΚΩΝ, ΟΙΚΟΝΟΜΙΚΩΝ ΚΑΙ
ΚΟΙΝΩΝΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΔΙΟΙΚΗΣΗΣ ΕΠΙΧΕΙΡΗΣΕΩΝ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ:
ΔΙΟΙΚΗΣΗ ΕΠΙΧΕΙΡΗΣΕΩΝ – MBA,
ΚΑΤΕΥΘΥΝΣΗ ΨΗΦΙΑΚΟ ΕΠΙΧΕΙΡΗΝ**

Μεταπτυχιακή Διπλωματική Εργασία

Τίτλος εργασίας

Εκτίμηση της επιτυχίας online προϊόντος με
Δείκτες Απόδοσης (KPIs)
χρησιμοποιώντας Μηχανική Μάθηση

Συγγραφείς

Ελένη Ταγκούτα
Παναγιώτης – Νικόλαος Ψυχάρης

AM:
MBA20046
MBA20049

Επιβλέπων:

Θεόδωρος Αναγνωστόπουλος

Αθήνα, Σεπτέμβριος 2022



UNIVERSITY OF WEST ATTICA
SCHOOL OF BUSINESS ADMINISTRATION,
ECONOMICS AND SOCIAL SCIENCES
DEPARTMENT OF BUSINESS ADMINISTRATION
POSTGRADUATE PROGRAM (MBA) – IN DIGITAL BUSINESS

Diploma Thesis

Title

Predicting success for web product
through KPIs
using Machine Learning

Student name and surname:

Eleni Tagkouta
Panagiotis – Nikolaos Psycharis

Registration Number:

MBA20046
MBA20049

Supervisor name and surname:

Theodoros Anagnostopoulos

Athens, September 2022



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ
ΣΧΟΛΗ ΔΙΟΙΚΗΤΙΚΩΝ, ΟΙΚΟΝΟΜΙΚΩΝ ΚΑΙ
ΚΟΙΝΩΝΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΔΙΟΙΚΗΣΗΣ ΕΠΙΧΕΙΡΗΣΕΩΝ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ:
ΔΙΟΙΚΗΣΗ ΕΠΙΧΕΙΡΗΣΕΩΝ – MBA,
ΚΑΤΕΥΘΥΝΣΗ ΨΗΦΙΑΚΟ ΕΠΙΧΕΙΡΗΝ**

**Τίτλος εργασίας
Εκτίμηση της επιτυχίας online προϊόντος με
Δείκτες Απόδοσης (KPIs)
χρησιμοποιώντας Μηχανική Μάθηση**

Μέλη Εξεταστικής Επιτροπής συμπεριλαμβανομένου και του Εισηγητή

Η μεταπτυχιακή διπλωματική εργασία εξετάστηκε επιτυχώς από την κάτωθι Εξεταστική Επιτροπή:

A/a	ΟΝΟΜΑ ΕΠΩΝΥΜΟ	ΒΑΘΜΙΑΔΑ/ΙΔΙΟΤΗΤΑ	ΨΗΦΙΑΚΗ ΥΠΟΓΡΑΦΗ
1.	Θεόδωρος Αναγνωστόπουλος	ΕΔΠΠ	
2.	Ιωάννης Σαλμόν	Αναπληρωτής Καθηγητής	
3.	Χρήστος Κυτάγιας	Επίκουρος Καθηγητής	

ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

Η κάτωθι υπογεγραμμένη Ελένη Ταγκούτα του Κωνσταντίνου, με αριθμό μητρώου MBA20046 φοιτήτρια του Προγράμματος Μεταπτυχιακών Σπουδών Διοίκηση Επιχειρήσεων (MBA) – Ψηφιακό Επιχειρήν του Τμήματος Διοίκησης Επιχειρήσεων της Σχολής Διοικητικών, Οικονομικών και Κοινωνικών Σπουδών του Πανεπιστημίου Δυτικής Αττικής, δηλώνω ότι:

«Είμαι συγγραφέας αυτής της μεταπτυχιακής εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της, είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, οι όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε ακριβώς είτε παραφρασμένες, αναφέρονται στο σύνολό τους, με πλήρη αναφορά στους συγγραφείς, τον εκδοτικό οίκο ή το περιοδικό, συμπεριλαμβανομένων και των πηγών που ενδεχομένως χρησιμοποιήθηκαν από το διαδίκτυο. Επίσης, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από μένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου, όσο και του Ιδρύματος.

Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του πτυχίου μου».

**Επιθυμώ την απαγόρευση πρόσβασης στο πλήρες κείμενο της εργασίας μου μέχρι
και έπειτα από αίτηση μου στη Βιβλιοθήκη και έγκριση του επιβλέποντα καθηγητή.*

Η Δηλούσα
Ελένη Ταγκούτα

*** Ονοματεπώνυμο /Ιδιότητα**

Ελένη Ταγκούτα / Μεταπτυχιακή Φοιτήτρια

Ψηφιακή Υπογραφή Επιβλέποντα
(Υπογραφή)

[CE%B8%CE%B5%CF%84%CE%B7%CF%81%CE%B9%CC%81%CE%BF%CF%85_final.pdf](#)

ΠΕΡΙΕΧΟΜΕΝΑ

1.	ΕΙΣΑΓΩΓΗ	7
2.	ΔΕΔΟΜΕΝΑ.....	13
3.	ΜΕΘΟΔΟΛΟΓΙΑ	16
3.1	Περιγραφή μεθοδολογίας.....	16
3.2	Περιγραφή Ερωτηματολογίων	17
3.3	Τρόπος επεξεργασίας ακατέργαστων δεδομένων	18
3.4	Περιγραφή Δεικτών Απόδοσης.....	19
3.5	Μετατροπή δεδομένων.....	21
3.5.1.	Εξαρτημένη μεταβλητή – Κωδικοποίηση	21
3.5.2	Ανεξάρτητες μεταβλητές - Κανονικοποίηση	21
3.6	Διαχωρισμός δεδομένων σε δεδομένα train – test	22
3.7	Περιγραφή του μοντέλου Νευρωνικού δικτύου.....	23
3.8	Confusion matrix.....	30
4.	ΑΠΟΤΕΛΕΣΜΑΤΑ.....	31
5.	ΣΥΖΗΤΗΣΗ.....	35
6.	ΣΥΜΠΕΡΑΣΜΑΤΑ	37
7.	ΒΙΒΛΙΟΓΡΑΦΙΑ	39
8.	ΠΑΡΑΡΤΗΜΑ.....	41

Επιτελική Σύνοψη

Η σημαντικότητα της έρευνας σχετίζεται με την συμβολή της μηχανικής μάθησης στο σχεδιασμό μιας στρατηγικής για βέλτιστη λήψη αποφάσεων και κατ' επέκταση του καθορισμού των ελάχιστων δυνατών λειτουργιών (features) που θα φέρει το προϊόν μιας ηλεκτρονικής πλατφόρμας από μια εταιρία τεχνολογίας, η οποία θα παρέχει χρηματοοικονομικές υπηρεσίες ώστε να είναι βιώσιμο (Minimum Viable Product - MVP), με σκοπό την σταδιακή ανάπτυξη και τελειοποίησή του. Για τις ανάγκες της ερευνητικής εργασίας ερευνάται το ενδεχόμενο δημιουργίας μιας ηλεκτρονικής πλατφόρμας από εταιρία τεχνολογίας, η οποία θα παρέχει χρηματοοικονομικές υπηρεσίες. Πιο συγκεκριμένα, θα συγκρίνει κατηγορίες δανείων μεταξύ των τραπεζών της Ελλάδας και θα παρέχει εξατομικευμένες προτάσεις κατόπιν λήψης πληροφοριών από τους ενδιαφερόμενους χρήστες με στόχο να προταθούν τα καταλληλότερα πακέτα για τις ανάγκες τους. Η συλλογή των απαραίτητων δεδομένων πραγματοποιήθηκε με χρήση ερωτηματολογίου όπου το δείγμα πάρθηκε με τυχαία δειγματοληψία και έγινε προσπάθεια να είναι όσο το δυνατόν μεγαλύτερο ώστε να είναι πιο αντιπροσωπευτικό. Γενικά, το ερωτηματολόγιο αποτελείται από τρεις ομάδες ερωτήσεων. Πρώτα τα δημογραφικά, δεύτερον ερωτήσεις που σχετίζονται με το αντικείμενο των υπηρεσιών που θα παρέχει η πλατφόρμα και τρίτον ερωτήσεις πάνω στις λειτουργίες που πρόκειται να παρέχει η πλατφόρμα. Ο σκοπός αυτής της έρευνας είναι να εκτιμήσει την επιτυχία της πλατφόρμας, χρησιμοποιώντας τις απαντήσεις των ερωτηθέντων και φτιάχνοντας στην συνέχεια ένα μοντέλο μηχανικής μάθησης. Για την εκτίμηση της επιτυχίας της πλατφόρμας θα χρησιμοποιηθούν Δείκτες Απόδοσης (Key Performance Indicators – KPIs) που βασίζονται στα δεδομένα που συλλέχθηκαν από τα ερωτηματολόγια, τα οποία στην συνέχεια εκτιμήθηκαν με την χρήση μηχανικής μάθησης. Πιο συγκεκριμένα, τα KPIs είναι δείκτες που χρησιμοποιούνται από οργανισμούς για τη μέτρηση, τη διαχείριση και τη σύγκριση της απόδοσης. Οι Δείκτες Απόδοσης μπορούν να οριστούν ως οι φυσικές αξίες που χρησιμοποιούνται για τη μέτρηση, τη σύγκριση και τη διαχείριση της συνολικής οργανωτικής απόδοσης. Επίσης, οι Δείκτες Απόδοσης μπορεί να περιλαμβάνουν την ποιότητα, τα χρηματοοικονομικά, την αξιοπιστία παράδοσης, την ικανοποίηση, την ασφάλεια, το περιβάλλον/κοινότητα και την μάθηση και ανάπτυξη. Έτσι, πραγματοποιήθηκε η εκτίμηση ενός KPI που αφορά την επιτυχία της πλατφόρμας με την συμβολή άλλων πέντε KPIs. Το μοντέλο που εφαρμόστηκε για την εκτίμηση του KPI είναι το Νευρωνικό Δίκτυο. Η αποτελεσματικότητα του μοντέλου προέκυψε 91,83%, και έτσι εκτιμάται ότι η ηλεκτρονική πλατφόρμα θα έχει μεγάλη επιτυχία (7/10 ανθρώπους που θα

ακούσουν για αυτή την online πλατφόρμα θα την χρησιμοποιήσουν) με 91,83% εγκυρότητα.

Λέξεις κλειδιά: Μηχανική Μάθηση, Νευρωνικά Δίκτυα, Επιχειρησιακό πλάνο, Δείκτες Απόδοσης – KPIs, Λήψη Αποφάσεων

1. ΕΙΣΑΓΩΓΗ

Η επιχειρηματική στρατηγική, αποτελεί ένα τεκμηριωμένο σχέδιο σχετικά με τον τρόπο με τον οποίο ένας οργανισμός επιδιώκει να επιτύχει τους στόχους του. Μια επιχειρηματική στρατηγική περιέχει μια σειρά από βασικές αρχές που περιγράφουν τον τρόπο με τον οποίο μια εταιρεία θα προχωρήσει στην επίτευξη αυτών των στόχων (Kitsios & Kamariotou, 2019). Για παράδειγμα, εξηγεί πώς θα αντιμετωπίσει τους ανταγωνιστές της, εξετάζει τις ανάγκες και τις προσδοκίες των πελατών και εξετάζει τη μακροπρόθεσμη ανάπτυξη και βιωσιμότητα του οργανισμού της. Ο λόγος για τον οποίο η ύπαρξη μιας στρατηγικής είναι τόσο σημαντική είναι επειδή δίνει χρόνο στις επιχειρήσεις να αποκτήσουν μια αίσθηση του πώς αποδίδουν, ποιες είναι οι δυνατότητές τους και αν αυτές οι δυνατότητες είναι σε θέση να τις βοηθήσουν να αναπτυχθούν.

Ένα εννοιολογικό πλαίσιο που μπορεί να λειτουργήσει ως συνολικό μοντέλο διαμόρφωσης στρατηγικής είναι αυτό των επιχειρηματικών ιδεών (Normann, 1975). Μια επιχειρηματική ιδέα θα μπορούσε να θεωρηθεί ως μια δήλωση που καθορίζει τον λόγο ύπαρξης μιας επιχείρησης, απαντώντας σε τρία βασικά ερωτήματα: Για ποιον, τι και πώς; Οι απαντήσεις σε αυτά τα τρία ερωτήματα περιλαμβάνουν μια περιγραφή της συγκεκριμένης αγοράς, του συστήματος προϊόντων που παρέχονται σε αυτή την αγορά και, τέλος, του συστήματος πόρων, ικανοτήτων και δομών που χρησιμοποιούνται για να πραγματοποιηθούν όλα αυτά.

Στην ψηφιακή εποχή, η δημιουργία επιχειρήσεων που έχουν την μορφή ιστοσελίδων για την παροχή υπηρεσιών παρατηρείται συνεχώς με νέες επιχειρήσεις να ακολουθούν αυτό τον τρόπο λειτουργίας από το ξεκίνημα τους. Κεντρικό ρόλο σε αυτό το φαινόμενο διαθέτει η αυξανόμενη ευφυΐα των εφαρμογών ιστού καθώς και η φορητότητά τους. Παρ' όλα αυτά, είναι περισσότερο αντιπροσωπευτικό του τρόπου με τον οποίο εφαρμόζεται η πληροφορική, σε αντίθεση με μια τεχνολογική επανάσταση, κυρίως επειδή πολλά από τα στοιχεία που αποτελούν τα θεμέλια του Ιστού 2.0 (Web 2.0) υπάρχουν ήδη από την αρχή της δημιουργίας του παγκόσμιου ιστού (world wide web) (Magumba, 2016).

Η τεχνολογία του Διαδικτύου έχει αλλάξει σημαντικά τους τρόπους με τους οποίους οι επιχειρήσεις συνεργάζονται και ανταγωνίζονται. Οι εταιρείες χρησιμοποιούν ολοένα και περισσότερο εφαρμογές ηλεκτρονικών επιχειρήσεων, όπως ηλεκτρονικές δημοπρασίες, ηλεκτρονικούς καταλόγους και εφαρμογές διαχείρισης πελατειακών σχέσεων, για την βελτιστοποίηση των επιχειρηματικών τους διαδικασιών σε ολόκληρη την αλυσίδα εφοδιασμού (Bakker, Zheng, & Knight, 2008). Οι Ηλεκτρονικές Επιχειρήσεις μπορούν να οριστούν ως συστήματα πληροφοριών για την απόκτηση, τη διεκπεραίωση, και τη μετάδοση πληροφοριών για την αποτελεσματικότερη λήψη αποφάσεων, σε σχέση με τα

ανταγωνιστικά πρότυπα. Το ηλεκτρονικό επιχειρείν αποτελεί μέρος ενός ευρύτερου οικονομικού πλαισίου που είναι υπεύθυνο για ριζικές μεταμορφώσεις στις επιχειρήσεις και περιλαμβάνει ψηφιακά δίκτυα και υποδομές επικοινωνίας (Borges, Hoppen, & Luce, 2009).

Οι επιχειρήσεις έχουν ξεκινήσει να χρησιμοποιούν σημαντικά το Διαδίκτυο και τις άλλες τεχνολογίες για να στηρίξουν τις εγχώριες και διεθνείς δραστηριότητές τους. Λόγω του αυξανόμενου αριθμού ηλεκτρονικών επιχειρήσεων, είναι όλο και πιο δύσκολο να μπορεί κανείς να διακριθεί στο Διαδίκτυο. Επομένως, ειδικά στο ηλεκτρονικό επιχειρείν είναι σημαντικό να υπάρχει η κατάλληλη επιχειρηματική στρατηγική διότι χωρίς αυτήν δεν μπορεί να υλοποιηθεί ούτε η καλύτερη ιδέα. Τα λάθη που γίνονται συχνότερα κατά την ανάπτυξη στρατηγικών για το ηλεκτρονικό επιχειρείν περιλαμβάνουν λανθασμένες προβλέψεις και αναλύσεις που προκύπτουν από την αποτυχία των δημιουργών της επιχείρησης να ελέγξουν αν η ιδέα τους θα προσελκύσει το ενδιαφέρον των πελατών και λανθασμένη εκτίμηση των οικονομικών χωρίς να λαμβάνονται υπόψη πρόσθετες πηγές χρηματοδότησης που θα επέτρεπαν στην επιχείρηση να αναπτυχθεί κατά την αρχική περίοδο, όταν δεν προσελκύει επαρκές ενδιαφέρον πελατών (Brzozowska, 2015).

Επειδή το περιβάλλον της αγοράς μπορεί να αλλάξει ταχύτερα από ό,τι τις προηγούμενες δεκαετίες, αρκετοί οργανισμοί έχουν υιοθετήσει αναδυόμενες τεχνολογίες που έχουν σχεδιαστεί για να αποκτήσουν υψηλές επιδόσεις και ανταγωνιστικό πλεονέκτημα. Μεταξύ αυτών των εξελίξεων, η Τεχνητή Νοημοσύνη (TN) κατέχει κομβική θέση (Ballog, 2020) και έχει προσελκύσει την προσοχή τόσο των ερευνητών όσο και του βιομηχανικού τομέα. Η τεχνητή νοημοσύνη αναφέρεται ως η ικανότητα μιας μηχανής να μαθαίνει από την εμπειρία, να προσαρμόζεται σε νέες καταστάσεις και να υλοποιεί εργασίες που μοιάζουν με τις ανθρώπινες (Duan, 2019). Σύμφωνα με τον (Duan, 2019), η τεχνητή νοημοσύνη θα μπορούσε πλέον να θεωρηθεί η οντότητα με τις σημαντικότερες δυνατότητες καινοτομίας και αποτελεί θεμελιώδη τεχνολογία πολλαπλών χρήσεων στον τομέα, ιδίως όσον αφορά τα εργαλεία μηχανικής μάθησης (Kitsios, 2021).

Την τελευταία δεκαετία, ο τεράστιος όγκος δεδομένων που συλλέγονται σε διάφορες μορφές είναι ταχύτερος από ποτέ. Αυτό επέβαλε την εμφάνιση νέων τεχνολογιών, με αποκορύφωμα την επιτάχυνση των τεχνολογικών εξελίξεων που ενσωματώνουν επίσης τις υπολογιστικές δυνατότητες επεξεργασίας, καθώς και την εξέλιξη νέων μεθόδων τεχνητής νοημοσύνης (Božič & Dimovski, 2019). Με αυτές τις εξελίξεις, οι επιχειρήσεις έχουν τη δυνατότητα να επεξεργάζονται τεράστιες ποσότητες δεδομένων με χρήση τεχνητής νοημοσύνης χρησιμοποιώντας τα αποτελέσματα για να διευρύνουν τον σκοπό τους για νέες αγορές, προϊόντα και υπηρεσίες (Sujata & Aniket).

Η εκθετική αύξηση στις διαδικτυακές πλατφόρμες που έχει παρουσιαστεί τα τελευταία χρόνια είχε ως αποτέλεσμα κάποιες από αυτές να αξιοποιούν τεχνολογίες μηχανικής μάθησης για την βελτίωση των υπηρεσιών που παρέχουν αλλά και για την στήριξη στη διαδικασία λήψης αποφάσεων στα άτομα που τις χρησιμοποιούν. Η πληθώρα υπηρεσιών που επιτρέπει η χρήση μηχανικής μάθησης δημιουργεί νέες ευκαιρίες για πολλές εταιρίες τεχνολογίας που συμβάλλει στον μεγάλο αριθμό αύξησης αυτού του τύπου ιστοσελίδων στο διαδίκτυο όπως φαίνεται και στο άρθρο του Whig (Whig, 2019). Συγκεκριμένα, αναφέρεται ότι η τεχνητή νοημοσύνη αποτελεί μέρος πολλών επιχειρήσεων σήμερα, διευκολύνει τις επιχειρηματικές δραστηριότητες, αυξάνει την παραγωγικότητα και προσφέρει μια ποικιλία τρόπων για την επιτάχυνση των διαδικασιών επικοινωνίας. Επίσης, η αυτοματοποίηση μέσω συστημάτων AI, εκτελεί πολλές από τις προηγούμενες εργασίες που πραγματοποιούνταν από εργαζόμενους και έχει ως αποτέλεσμα μια επιχείρηση να μειώσει τα έξοδά της, να σημειώσει σημαντική εξοικονόμηση χρόνου και επίσης σταδιακή αύξηση των κερδών.

Εκτός από την μηχανική μάθηση, υποπεδίο της τεχνητής νοημοσύνης αποτελεί και η Βαθιά Μάθηση (Deep Learning), οι εφαρμογές της οποίας έχουν πετύχει εντυπωσιακά αποτελέσματα σε διάφορους τομείς μέσα στα λίγα σχετικά χρόνια που έχει εμφανιστεί. Η Βαθιά Μάθηση στηρίζεται σε νευρωνικά δίκτυα (neural nets) που αποτελούνται από μεγάλο αριθμό απλών μονάδων, τους νευρώνες, οι οποίοι ανταλλάσσουν σήματα μεταξύ τους και συγκροτούν ένα μοντέλο που χρησιμοποιείται για την επίλυση υπολογιστικών προβλημάτων (Goodfellow, 2016). Λόγω της ικανότητας της να επεξεργάζεται τις πληροφορίες με παρόμοιο τρόπο με τον οποίο οι άνθρωποι αποκτούν ορισμένους τύπους γνώσεων και της υψηλότερης ακρίβειας που προσφέρει σε σχέση με τις συμβατικές μεθόδους της μηχανικής μάθησης, επιλέχθηκε η χρήση της σε αυτή την ερευνητική εργασία. Πιο αναλυτικά, η εργασία ασχολείται με την εκτίμηση της επιτυχίας online προϊόντος (πλατφόρμα χρηματοοικονομικών υπηρεσιών) με δείκτες απόδοσης (KPIs) κάνοντας χρήση νευρωνικών δικτύων όπου συλλέχθηκαν δεδομένα μέσω ερωτηματολογίων με πρωταρχικό στόχο να υπολογιστεί πόσο πιθανό είναι να χρησιμοποιήσει κάποιος την πλατφόρμα και δεύτερον, για την εύρεση των σημαντικότερων χαρακτηριστικών/λειτουργιών που πρέπει να διαθέτει για να είναι πιο επιθυμητή από στους υποψήφιους χρήστες.

Το παράδειγμα της ερευνητικής εργασίας στο οποίο εφαρμόστηκε το νευρωνικό δίκτυο αφορά μια πλατφόρμα χρηματοοικονομικών υπηρεσιών που βρίσκεται σε στάδιο σχεδιασμού και ανάλυσης ώστε να αποφασιστεί ποια θα είναι τα βασικά χαρακτηριστικά που θα διαθέτει στο ξεκίνημά της ως Ελάχιστο Βιώσιμο Προϊόν (Minimum Viable Product - MVP). Πιο συγκεκριμένα, η πλατφόρμα θα συγκρίνει κατηγορίες δανείων μεταξύ των

τραπεζών της Ελλάδας και θα παρέχει εξατομικευμένες προτάσεις κατόπιν λήψης πληροφοριών από τους ενδιαφερόμενους χρήστες με στόχο να προταθούν τα καταλληλότερα πακέτα για τις ανάγκες τους.

Η μέθοδος Βαθιάς Μάθησης με νευρωνικά δίκτυα που χρησιμοποιήθηκε σε αυτή την εργασία ακολουθεί την προσέγγιση της μάθησης με επίβλεψη (Supervised Learning). Σε αυτή την προσέγγιση, ένας αλγόριθμος υπολογιστή εκπαιδεύεται σε δεδομένα εισόδου που έχουν επισημανθεί για μια συγκεκριμένη έξοδο με στόχο να ταξινομήσει δεδομένα ή να προβλέψει αποτελέσματα με ακρίβεια. Συγκεκριμένα, εάν η τιμή εξόδου υπερβαίνει ένα δεδομένο όριο, "πυροδοτεί" ή ενεργοποιεί τον κόμβο, μεταφέροντας δεδομένα στο επόμενο επίπεδο του δικτύου. Τα νευρωνικά δίκτυα μαθαίνουν αυτή τη συνάρτηση αντιστοίχισης μέσω μάθησης με επίβλεψη, προσαρμοζόμενα με βάση τη συνάρτηση απωλειών (loss function) μέσω της διαδικασίας Gradient Descent (IBM, 2020). Όταν η συνάρτηση κόστους (cost function) είναι στο μηδέν ή κοντά στο μηδέν, μπορούμε να είμαστε σίγουροι για την ακρίβεια του μοντέλου να δώσει τη σωστή απάντηση.

Η μάθηση με επίβλεψη αποτελεί την πιο συνηθισμένη από τις μεθόδους μάθησης στο αντικείμενο της τεχνητής νοημοσύνης. Η μη επιβλεπόμενη μάθηση (Unsupervised Learning) που χρησιμοποιεί μη επισημασμένα δεδομένα και η ενισχυτική μάθηση (Reinforcement Learning) όπου ένας πράκτορας λαμβάνει πληροφορίες για το περιβάλλον του και μαθαίνει να επιλέγει ενέργειες που μεγιστοποιούν κάποια ανταμοιβή αποτελούν άλλες μεθόδους που χρησιμοποιούνται όλο και συχνότερα μεταξύ άλλων (Chollet, 2018). Η μάθηση με επίβλεψη μπορεί να διαχωριστεί σε δύο τύπους προβλημάτων, την ταξινόμηση (που χρησιμοποιήθηκε σε αυτή την εργασία) και την παλινδρόμηση. Η ταξινόμηση χρησιμοποιεί έναν αλγόριθμο για την ακριβή ανάθεση δεδομένων δοκιμής σε συγκεκριμένες κατηγορίες. Αναγνωρίζει συγκεκριμένες οντότητες στο σύνολο δεδομένων και προσπαθεί να βγάλει κάποια συμπεράσματα σχετικά με τον τρόπο με τον οποίο αυτές οι οντότητες θα πρέπει να επισημανθούν ή να οριστούν. Από την άλλη πλευρά, η παλινδρόμηση χρησιμοποιείται για την κατανόηση της σχέσης μεταξύ εξαρτημένων και ανεξάρτητων μεταβλητών. Χρησιμοποιείται συνήθως για την πραγματοποίηση προβλέψεων, όπως για τα έσοδα από τις πωλήσεις μιας συγκεκριμένης επιχείρησης (IBM, 2020).

Όσον αφορά την συλλογή των δεδομένων, πραγματοποιήθηκε με χρήση ερωτηματολογίου όπου το δείγμα συλλέχθηκε με τυχαία δειγματοληψία και έγινε προσπάθεια να είναι όσο το δυνατόν μεγαλύτερο ώστε να είναι πιο αντιπροσωπευτικό. Γενικά, το ερωτηματολόγιο που δημιουργήθηκε διέθετε τρεις (3) ομάδες ερωτήσεων. Πρώτα τα δημογραφικά, δεύτερον ερωτήσεις που σχετίζονται με το αντικείμενο των

υπηρεσιών που θα παρέχει η πλατφόρμα και τρίτον ερωτήσεις πάνω στις λειτουργίες που πρόκειται να παρέχει η πλατφόρμα.

Ο σκοπός αυτής της έρευνας είναι να εκτιμήσει την επιτυχία της πλατφόρμας, χρησιμοποιώντας τις απαντήσεις των ερωτηθέντων και φτιάχνοντας στην συνέχεια ένα μοντέλο μηχανικής μάθησης χρησιμοποιώντας ένα νευρωνικό δίκτυο. Για την εκτίμηση της επιτυχίας της πλατφόρμας χρησιμοποιήθηκαν Δείκτες Απόδοσης (Key Performance Indicators – KPIs) που βασίζονται στα δεδομένα που συλλέχθηκαν από τα ερωτηματολόγια, τα οποία στην συνέχεια εκτιμήθηκαν με την χρήση μηχανικής μάθησης. Πιο συγκεκριμένα, τα KPIs είναι δείκτες που χρησιμοποιούνται από οργανισμούς για τη μέτρηση, τη διαχείριση και τη σύγκριση της απόδοσης. Οι δείκτες απόδοσης μπορούν να οριστούν ως οι φυσικές αξίες που χρησιμοποιούνται για τη μέτρηση, τη σύγκριση και τη διαχείριση της συνολικής οργανωτικής απόδοσης (Awan, 2014).

Η σωστή οργάνωση του ερωτηματολογίου ήταν σημαντική για την συγκέντρωση των κατάλληλων δεδομένων γι' αυτό αξιοποιήθηκε το βιβλίο της Davino και του Fabbris, Survey Data Collection and Integration (2013), όπου εξετάζονται ζητήματα τόσο στην θεωρία όσο και στις εφαρμογές της έρευνας. Πιο συγκεκριμένα, το βιβλίο αυτό παρέχει λύσεις για την αντιμετώπιση προβλημάτων συλλογής και ανάλυσης δεδομένων, για τον σχεδιασμό δειγμάτων και ερωτηματολογίων αλλά και για την στατιστική εκτίμηση. Συνεπώς, σε αυτό το βιβλίο, εστιάζει το κομμάτι της έρευνας που αφορά την δημιουργία των ερωτηματολογίων, την δειγματοληψία και την συλλογή δεδομένων για την δημιουργία των KPIs.

Το βιβλίο των Provost και Fawcett (2013), με τίτλο Data science for business χρησιμοποιήθηκε για την ερμηνεία των αναλύσεων που έγιναν στα δεδομένα που συλλέχθηκαν και πώς μπορούν να ενισχύουν μια επιχείρηση που προωθεί μια πλατφόρμα αυτού του είδους. Επιπροσθέτως, στο συγκεκριμένο βιβλίο αναφέρονται έννοιες που αποτελούν την βάση της ανάλυσης επιχειρησιακών προβλημάτων, με επίκεντρο τα δεδομένα, την δημιουργία και αξιολόγηση λύσεων, αλλά και γενικών στρατηγικών και προτάσεων. Στοχεύει κυρίως στο να κάνει κατανοητή την επιστήμη των δεδομένων στους αναγνώστες, μέσα από έννοιες και μεγάλη ποικιλία μεθόδων και τεχνικών επιχειρησιακής ανάλυσης. Οι έννοιες αυτές αφορούν τρεις γενικούς τύπους, η πρώτη εστιάζει στον τρόπο συμβολής της επιστήμης δεδομένων στο ανταγωνιστικό τοπίο, η δεύτερη στον εντοπισμό κατάλληλων δεδομένων και κατάλληλων μεθόδων και η τρίτη στην εξαγωγή γνώσης από τα δεδομένα. Το βιβλίο αυτό, βοήθησε στο θεωρητικό κομμάτι της έρευνας, σχετικά με τις έννοιες που θα αναφερθούν για την επιστήμη δεδομένων.

Ιδιαίτερα σημαντικό μέρος σε διαδικασίες που περιλαμβάνουν ανάλυση δεδομένων είναι η οπτικοποίηση, η οποία μπορεί να δημιουργήσει μια σαφέστερη και πιο

ολοκληρωμένη κατανόηση των δεδομένων. Στο βιβλίο του Witney, Data Insights (2012), διερευνώνται τρόποι με τους οποίους η οπτικοποίηση δεδομένων μπορεί να κάνει τα δεδομένα κατανοητά και χρήσιμα. Εξετάζονται επίσης οι βασικές αρχές, η προέλευση και ο σκοπός των βασικών δομικών στοιχείων που χρησιμοποιούνται στην οπτικοποίηση δεδομένων. Το βιβλίο αυτό, βοήθησε στην καλύτερη δημιουργία γραφημάτων για την παρουσίαση των αποτελεσμάτων, που αφορούν τόσο την εκτίμηση των KPIs αλλά και των υπόλοιπων αποτελεσμάτων.

Προχωρώντας στον ρόλο που είχαν οι βασικοί δείκτες απόδοσης (KPIs) στην εργασία, σύμφωνα με το άρθρο των Villazón και άλλοι, 2020, αναφέρουν τον προσδιορισμό των βασικών δεικτών απόδοσης (KPIs) σε οργανισμούς, που βασίζονται σε έργα, με βάση τις οργανωτικές και λειτουργικές τους ανάγκες. Η έρευνά τους εστιάζει κυρίως στην κατηγοριοποίηση των KPIs μέσω μιας ποιοτικής προσέγγισης για την επιτυχία ενός έργου. Παρουσιάζεται συγκεκριμένα μια μελέτη περίπτωσης μιας ανάλυσης ενός οργανισμού Έρευνας και Ανάπτυξης (Research & Development – (R&D)). Τονίζεται η σημασία για τον επαναπροσδιορισμό των KPIs εξαιτίας των συνεχών αλλαγών στην αγορά ή τις μεθοδολογίες της έρευνας, ώστε να διασφαλιστεί ότι είναι τα καταλληλότερα για το υπάρχον περιβάλλον του οργανισμού. Η έρευνα αυτή, είχε βασικό ρόλο στη δημιουργία των κατάλληλων KPIs, ώστε να διαμορφώνονται βάσει των στρατηγικών και επιχειρηματικών στόχων καθώς και σύμφωνα με την εταιρική κουλτούρα της επιχείρησης που ετοιμάζει την συγκεκριμένη πλατφόρμα.

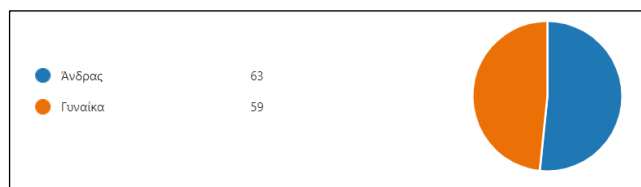
Συνοψίζοντας, πρωταρχικός σκοπός της ερευνητικής εργασίας είναι η εύρεση της πιθανότητας επιτυχίας της πλατφόρμας με χρήση ενός νευρωνικού δικτύου χρησιμοποιώντας τα δεδομένα που έχουν συλλεχθεί από τα ερωτηματολόγια. Επιπροσθέτως, θα εντοπίσουμε τα κατάλληλα χαρακτηριστικά και λειτουργίες που πρέπει να υποστηρίζει η πλατφόρμα στην πρώτη της έκδοση ως Ελάχιστο Βιώσιμο Προϊόν, κάνοντας ανάλυση των απαντήσεων από τις αντίστοιχες ερωτήσεις του ερωτηματολογίου που μοιράστηκε.

2. ΔΕΔΟΜΕΝΑ

Τα δεδομένα που χρησιμοποιήθηκαν για αυτή την έρευνα, αφορούν ανώνυμα προσωπικά δεδομένα σχετικά με τραπεζικά δάνεια. Πιο συγκεκριμένα, τα δεδομένα αφορούν δημογραφικά στοιχεία, πληροφορίες σχετικά με το ιστορικό των τραπεζικών δανείων που έχουν προμηθευτεί οι ερωτηθέντες στο παρελθόν αλλά και κάποια άλλα γενικά στοιχεία για την χρηματοοικονομική διαδικτυακή πλατφόρμα που θα δημιουργηθεί.

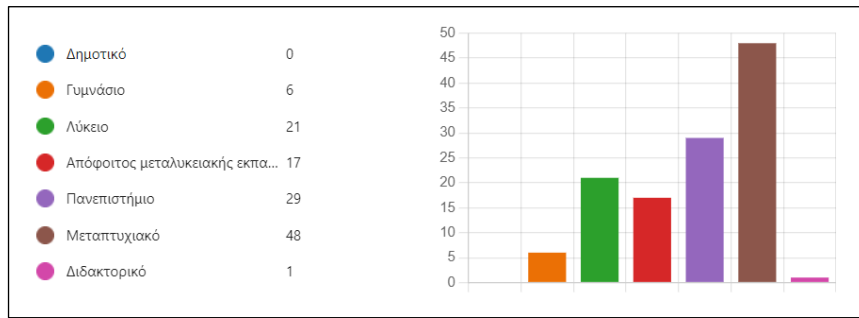
Ιδιαίτερα, τα δεδομένα λήφθηκαν με την χρήση ερωτηματολογίων, όπου το πλήθος των συμπληρωμένων ερωτηματολογίων είναι 122 και το καθένα από αυτά αποτελείται από 32 ερωτήσεις. Κάθε συμπληρωμένο ερωτηματολόγιο αφορά και μια εγγραφή, ενώ κάθε ερώτηση αφορά και μια διάσταση στον χώρο των μεταβλητών. Έτσι, για την ανάλυσή τους τα αρχικά ακατέργαστα δεδομένα τοποθετήθηκαν σε δομή ενός Πίνακα (122, 32) διαστάσεων.

Επίσης, το ερωτηματολόγιο είναι διαδικτυακό (online) και δημιουργήθηκε και διαμοιράστηκε στους ανθρώπους μέσω του Microsoft Forms, ένα εργαλείο που παρέχει δωρεάν την κατασκευή, το διαμοίρασμα και την συλλογή των δεδομένων. Όσον αφορά το δείγμα που πάρθηκε, παρακάτω παρουσιάζονται κάποια δημογραφικά στοιχεία του πληθυσμού αυτού, *Εικόνες 1 – 3* και *Πίνακας 1*.



Εικόνα 1: Κατηγορίες Φύλου δείγματος

Στην *Εικόνα 1*, όσον αφορά το Φύλο, παρατηρείται ότι το δείγμα της παρούσας εργασίας αποτελείται από σχεδόν ίσο αριθμό ανδρών (63/122) και των γυναικών (59/122), με τους άνδρες να υπερिशύουν ελάχιστα.



Εικόνα 2: Επίπεδο Εκπαίδευσης δείγματος

Σχετικά με το Επίπεδο Εκπαίδευσης του δείγματος, στην *Εικόνα 2* παρατηρείται ότι η πλειοψηφία των συμμετεχόντων είναι κάτοχος μεταπτυχιακού διπλώματος (48/122), ενώ οι λιγότεροι (0/122), δηλαδή κανένας, δεν έχει δηλώσει ότι δεν κατέχει Απολυτήριο Δημοτικού Σχολείου.



Εικόνα 3: Καθαρό μηνιαίο Εισόδημα (Ευρώ) του δείγματος

Σχετικά με το Μηνιαίο Καθαρό Εισόδημα (σε Ευρώ) στην *Εικόνα 3*, το μεγαλύτερο μέρος του δείγματος (35/122) λαμβάνει έναν μέτριο μισθό (1.001 – 1.500 ευρώ), ενώ το μικρότερος μέρος (3/122) λαμβάνει έναν αρκετά μεγάλο μισθό (2.001 ευρώ και άνω).

Πίνακας 1: Βαθμός Γνώσης Η/Υ του δείγματος

Μέσος όρος: 122 απαντήσεις	3.8
-----------------------------------	-----

Ακόμη, η γνώση της πλειοψηφίας του δείγματος όσον αφορά την γνώση τους σε Ηλεκτρονικούς Υπολογιστές, όπως φαίνεται και στον Πίνακα 1, είναι μεγάλη (3.8/5) με [1: καθόλου, 2: λίγο, 3: μέτρια, 4: πολύ, 5: πάρα πολύ].

Συνοψίζοντας, τα βασικά δημογραφικά χαρακτηριστικά του δείγματος όσον αφορά την πλειοψηφία του αφορούν ανθρώπους αρκετά μορφωμένους με μέτριο εισόδημα και αρκετά καλή γνώση Η/Υ.

3. ΜΕΘΟΔΟΛΟΓΙΑ

3.1 Περιγραφή μεθοδολογίας

Η μεθοδολογία που ακολουθήθηκε αφορά εμπειρική έρευνα, δηλαδή την σύνδεση της θεωρίας με την πραγματικότητα (Sauce και Matzel, 2017) και επομένως είναι Επαγωγική. Πιο συγκεκριμένα, η Επαγωγική έρευνα ξεκινάει από συγκεκριμένες παρατηρήσεις και καταλήγει σε ευρύτερες γενικεύσεις και θεωρίες και περιλαμβάνει έναν βαθμό αβεβαιότητας (Burney και Saleem, 2008). Ο Επαγωγικός συλλογισμός συχνά χρησιμοποιείται για να δημιουργήσει προβλέψεις (Sauce και Matzel, 2017), όπως η συγκεκριμένη έρευνα.

Επίσης, είναι μελέτη περίπτωσης (case study). Μια μελέτη περίπτωσης <<αφηγείται>> μια ενδιαφέρουσα ιστορία σχετικά με άτομα, οργανισμούς, διαδικασίες, γεγονότα κλπ. και εστιάζει σε πραγματικά φαινόμενα της ζωής τα οποία όμως δεν είναι σαφώς εμφανή. Τα φαινόμενα αυτά προσπαθεί να τα αναλύσει και να τα εξηγήσει. Ο σκοπός της είναι να ερευνήσει εις βάθος μέσω της ανάλυσης και να παρέχει περισσότερες αλλά και πιο ενδιαφέρουσες πληροφορίες. Η μελέτη περίπτωσης εφαρμόζεται και στις επιχειρήσεις με την ίδια λογική. Ο σχεδιασμός της μελέτης περίπτωσης συνδέει τα εμπειρικά δεδομένα με τα ερευνητικά ερωτήματα και καταλήγει σε συμπεράσματα (Norhayate και Rahim, 2015).

Όσον αφορά την συλλογή δεδομένων, πραγματοποιήθηκε μέσω ερωτηματολογίου, όπου για να δημιουργηθεί με τον κατάλληλο τρόπο πραγματοποιήθηκε μελέτη για την συγκεκριμένη έρευνα. Γενικά, ένα ερωτηματολόγιο θεωρείται ένα βοήθημα για την συλλογή των δεδομένων για μια ανάλυση. Για τον λόγο αυτό, στη σχεδιάσή του οι ερευνητές θα πρέπει να έχουν κατά νου το ευρύτερο πλαίσιο στο οποίο θα χρησιμοποιηθεί το ερωτηματολόγιο. Για παράδειγμα, τους στόχους της έρευνας, το είδος των ερωτήσεων που θα τεθούν, το πλήθος των ανθρώπων που θα απαντήσουν, ποιοι θα συμπεριληφθούν στο δείγμα της μελέτης, πως θα παραδοθεί, πως θα συμπληρωθεί κλπ. (Ekinici, 2015).

Στην συνέχεια, έγινε ποσοτική ανάλυση με την χρήση της μηχανικής μάθησης και συγκεκριμένα με την εφαρμογή της ενός μοντέλου Νευρωνικού Δικτύου και εξάχθηκαν αποτελέσματα για την εκτίμηση των KPIs, των λειτουργιών (features) που πρόκειται να εισαχθούν στο προϊόν. Τέλος, εξάχθηκαν συμπεράσματα για την έρευνα αλλά και κάποια σχόλια πάνω στο πεδίο της έρευνας όπως και τα μελλοντικά βήματά της.

3.2 Περιγραφή Ερωτηματολογίων

Δεδομένου ότι μια έρευνα στοχεύει στο να βγάλει συμπεράσματα για έναν πληθυσμό, για τον λόγο αυτό και σε αυτή την έρευνα έγινε η χρήση του ερωτηματολογίου ως εργαλείο για την συλλογή των δεδομένων, με σκοπό την ανάλυσή τους.

Ένα σημαντικό για την κατασκευή του ερωτηματολογίου που χρησιμοποιήθηκε σε αυτή την έρευνα, είναι το δείγμα (Sample). Υπάρχουν δύο βασικές κατηγορίες δειγματοληψίας, η Δειγματοληψία Πιθανοτήτων και η Δειγματοληψία Μη – Πιθανοτήτων. Η πρώτη περίπτωση αφορά την συλλογή ενός τυχαίου δείγματος, έτσι ώστε η κάθε μονάδα να έχει την ίδια πιθανότητα να επιλεγεί, με ένα κοινό κριτήριο επιλογής (Bryman, 2004). Η περίπτωση αυτή είναι μεν πιο επιθυμητή επειδή είναι πιο αντιπροσωπευτικό το δείγμα και μειώνει την πιθανότητα δειγματοληπτικού σφάλματος, αλλά από την άλλη στην πράξη είναι πολύ δύσκολο να παρθεί εξαιτίας της πολύ μικρής κλίμακας του διαθέσιμου πληθυσμού (Döngyi, 2007). Η δεύτερη περίπτωση δειγματοληψίας αποτελείται από κάποιες κατηγορίες, όπως για παράδειγμα η Δειγματοληψία Χιονοστιβάδας που χρησιμοποιείται και στην εκάστοτε έρευνα. Η χρήση αυτής της μεθοδολογίας επιλέχθηκε μεταξύ των υπολοίπων λόγω της ευκολίας της να χρησιμοποιηθεί με την χρήση των μέσων κοινωνικής δικτύωσης.

Στην Δειγματοληψία Χιονοστιβάδας, οι ερευνητές έρχονται σε επαφή με έναν μικρό αριθμό ατόμων, ο οποίος με την σειρά του τα προωθεί σε δικό του αριθμό ατόμων και ούτω κάθε εξής. Στην περίπτωση αυτή, το δείγμα δεν είναι τυχαίο, έχει όμως το πλεονέκτημα ότι η δειγματοληψία είναι γρήγορη, φθηνή και ένας είναι ένας βολικός τρόπος συλλογής δεδομένων (Young, 2015). Στην έρευνα αυτή, το ερωτηματολόγιο στάλθηκε σε έναν μικρό αριθμό ατόμων στα Μέσα Κοινωνικής Δικτύωσης (Facebook, Viber, Email) όπου με την σειρά τους τα άτομα αυτά τα προώθησαν σε επιπλέον άτομα και έτσι πάρθηκε το δείγμα με την χρήση της δειγματοληψίας της Χιονοστιβάδας.

Ένα ακόμα σημαντικό στοιχείο για την κατασκευή ενός ερωτηματολογίου είναι ο Σχεδιασμός της μορφής του (Format design). Η συνολική εμφάνιση ενός ερωτηματολογίου είναι ζωτικής σημασίας. Όσο πιο ξεκάθαρη διάταξη και δομή έχει τόσο μεγαλύτερη πιθανότητα έχει να συμπληρωθεί μέχρι το τέλος. Θα πρέπει να υπάρχει γενικά μια ισορροπία σε αυτό για να είναι το βέλτιστο. Όσον αφορά το μέγεθός του θα πρέπει να μην απαιτεί πάρα πολύ χρόνο για την συμπλήρωσή του. Επίσης, η γραμματοσειρά του θα πρέπει να είναι σε ένα μέγεθος φιλικό προς τον χρήστη για την ανάγνωσή του. Δεν θα πρέπει να υπάρχουν πάρα πολλές ερωτήσεις σχετικά με τα δημογραφικά στοιχεία που στην πραγματικότητα δεν ενδιαφέρουν τον ερευνητή. Επίσης, να έχει σαφής και απλές οδηγίες για την συμπλήρωσή του. Εάν δεν ληφθούν αυτά υπόψιν, υπάρχει πιθανότητα να αποτρέψει τους ανθρώπους να απαντήσουν. Γενικά, ένα ερωτηματολόγιο θα πρέπει να

δημιουργείται με γνώμονα την φύση της έρευνας, δηλαδή του προβλήματος που έχει τεθεί αρχικά και να ακολουθεί την κλασσική δομή/μορφή των ερωτηματολογίων που είναι διεθνώς αποδεκτή. Στο ερωτηματολόγιο αυτής της έρευνας, λήφθηκαν υπόψιν όλα τα παραπάνω με σκοπό την βέλτιστη λήψη των δεδομένων.

Πιο συγκεκριμένα, ο τύπος των ερωτήσεων μπορεί να είναι είτε <<κλειστές>> είτε <<ανοιχτές>>. Οι κλειστές έχουν σαφή και συγκεκριμένες απαντήσεις (μία ή πολλές). Αντίθετα, οι ανοιχτού τύπου δεν δίνουν ιδιαίτερες επιλογές και αφήνουν τους ανθρώπους να απαντήσουν με τον δικό τους τρόπο. Στην έρευνα αυτή, έγινε χρήση των κλειστών απαντήσεων, διότι έχουν το πλεονέκτημα της συλλογής των ποσοτικών δεδομένων, που είναι πιο εύκολο να κωδικοποιηθούν σε μεταβλητές αριθμητικά και να έτσι να αναλυθούν στατιστικά (Young, 2015). Επίσης, οι κλίμακες μέτρησης αυτών των μεταβλητών της έρευνας είναι δύο, οι Ονομαστικές (Nominal) και οι Τάξης (Ordinal), χωρίς την χρήση και των Συνεχών (Scale). Οι ονομαστικές αφορούν δεδομένα που είναι απλά ονόματα, δηλαδή κατηγορίες, οι οποίες δεν έχουν κάποια φυσική κατάταξη, π.χ. το φύλο. Από την άλλη, οι τάξης είναι και πάλι κατηγορίες οι οποίες όμως έχουν μια σαφή σειρά, π.χ. τα δεδομένα κλίμακας Likert από διαβάθμιση σε ‘Διαφωνώ πολύ’ έως ‘Συμφωνώ πολύ’. Ακόμη, οι Συνεχής, που δεν χρησιμοποιήθηκαν στην εκάστοτε έρευνα, είναι ποσοτικά δεδομένα που μετριούνται σε τιμές π.χ. η ηλικία (Mishra και άλλοι, 2018).

Συγκεκριμένα, όπως έχει αναφερθεί και σε προηγούμενο κεφάλαιο, το ερωτηματολόγιο αποτελείται από τρεις ομάδες ερωτήσεων. Πρώτα τα δημογραφικά, δεύτερον ερωτήσεις που σχετίζονται με το αντικείμενο των υπηρεσιών που θα παρέχει η πλατφόρμα και τρίτον ερωτήσεις πάνω στις λειτουργίες που πρόκειται να παρέχει η πλατφόρμα.

3.3 Τρόπος επεξεργασίας ακατέργαστων δεδομένων

Δεδομένου ότι οι απαντήσεις των ερωτήσεων είναι της μορφής των <<κλειστών>> ερωτήσεων, σε κάποιες από αυτές που οι απαντήσεις είναι της μορφής μικρού μήκους κειμένου ήταν απαραίτητη η κωδικοποίησή τους σε αριθμητικές τιμές. Ο σκοπός της ποσοτικοποίησής τους ήταν για να πραγματοποιηθεί η βέλτιστη ανάλυση των δεδομένων και μετέπειτα να δημιουργηθούν οι κατάλληλοι Δείκτες Απόδοσης (KPIs).

3.4 Περιγραφή Δεικτών Απόδοσης

Στην παρούσα έρευνα εκτιμήθηκε η επιτυχία μιας online μελλοντικής πλατφόρμας, χρησιμοποιώντας τις απαντήσεις των ερωτηθέντων και φτιάχνοντας στην συνέχεια ένα μοντέλο μηχανικής μάθησης. Για την δημιουργία του μοντέλου αυτού, δημιουργήθηκαν έξι Δείκτες Απόδοσης (KPIs). Έτσι, ο ένας Δείκτης Απόδοσης (KPI 1) χρησιμοποιήθηκε ως εξαρτημένη μεταβλητή y , και οι υπόλοιποι πέντε (KPI 2 έως KPI 6) ως ανεξάρτητες μεταβλητές X_1, \dots, X_N , με $i = 2$ και $N = 6$. Ακολουθεί ο Πίνακας 2, με τους Δείκτες Απόδοσης:

Πίνακας 2: Δείκτες Απόδοσης έρευνας - KPIs

	Δείκτης Απόδοσης – KPI
KPI 1 <i>εξαρτημένη μεταβλητή</i>	Ποσοστό % των εν δυνάμει χρηστών της πλατφόρμας
KPI 2 <i>ανεξάρτητη μεταβλητή</i>	Σύνολο χρηστών ιστοσελίδας τράπεζας
KPI 3 <i>ανεξάρτητη μεταβλητή</i>	Σύνολο ευχαριστημένων και μη πολιτών που έχουν χρησιμοποιήσει ιστοσελίδες τραπεζών
KPI 4 <i>ανεξάρτητη μεταβλητή</i>	Σύνολο ποσού δανείου
KPI 5 <i>ανεξάρτητη μεταβλητή</i>	Αναλογία επιπέδου εκπαίδευσης και της γνώση τους σε H/Y
KPI 6 <i>ανεξάρτητη μεταβλητή</i>	Σύνολο εργαζομένων ανά νοικοκυριό

Οι συγκεκριμένοι Δείκτες Απόδοσης επιλέχθηκαν να δημιουργηθούν, διότι μπορούν να δίνουν την πιο σχετική πληροφορία σε σχέση με άλλες πληροφορίες που πάρθηκαν από τα ερωτηματολόγια. Ιδιαίτερα, δεδομένου ότι η παρούσα έρευνα σχετίζεται με ένα online προϊόν για τραπεζικά δάνεια για τον λόγο αυτό, επιλέχθηκαν Δείκτες Απόδοσης που αφορούν ικανότητες των ερωτηθέντων για χρήση H/Y, χρήση ιστοσελίδων και κάποια δημογραφικά στοιχεία τους τα οποία παραπέμπουν σε χρήματα (π.χ. εισόδημα, εργασία κλπ.) τα οποία συνήθως ακολουθούν κάποιο μοτίβο ανάλογα την φύση του προβλήματος. Επίσης, αξίζει να αναφερθεί αυτοί Δείκτες Απόδοσης επιλέχθηκαν για να σχετίζονται με τους στόχους που έχει θέσει η επιχείρηση στο επιχειρηματικό της πλάνο, με σκοπό την αύξηση του κέρδους της. Παρακάτω ακολουθεί η περιγραφή των Δεικτών Απόδοσης:

- **Ποσοστό % των εν δυνάμει χρηστών της πλατφόρμας.**

Εδώ, ως εξαρτημένη μεταβλητή, είναι μια μετρική (measure) που στοχεύει στην εκτίμηση του Ποσοστού των ανθρώπων οι οποίοι θα χρησιμοποιήσουν την πλατφόρμα αυτή μετά την ενημέρωσή τους για την ύπαρξή της. Ως αποτέλεσμα, έχει την αύξηση του κέρδους της επιχείρησης με ταυτόχρονη μείωση του ρίσκου της, για την δημιουργία της πλατφόρμας που πρόκειται πραγματοποιηθεί.

- **Σύνολο χρηστών ιστοσελίδας τράπεζας.**

Εδώ, ως ανεξάρτητη μεταβλητή, είναι μια μετρική (measure) που στοχεύει στο Σύνολο εκείνων των ανθρώπων που χρησιμοποιούν ιστοσελίδες τραπεζών και μπορεί να δείξει ότι οι άνθρωποι αυτοί είναι πιο εξοικειωμένοι με την χρήση τους. Επομένως, είναι πιο πιθανοί χρήστες της πλατφόρμας.

- **Σύνολο ευχαριστημένων και μη πολιτών που έχουν χρησιμοποιήσει ιστοσελίδες τραπεζών.**

Εδώ, ως ανεξάρτητη μεταβλητή, είναι μια μετρική (measure) που στοχεύει στην εύρεση της ικανοποίησης των εν δυνάμει χρηστών της πλατφόρμας. Όσο πιο ικανοποιημένοι είναι οι χρήστες των ιστοσελίδων των τραπεζών, υπάρχει μεγάλη πιθανότητα να χρησιμοποιήσουν και την πλατφόρμα που πρόκειται να δημιουργηθεί.

- **Σύνολο ποσού δανείου.**

Εδώ, ως ανεξάρτητη μεταβλητή, είναι μια μετρική (measure) που στοχεύει στην εύρεση του Συνόλου των δανείων που έλαβαν οι ερωτηθέντες. Αυτό χρησιμοποιήθηκε ως Δείκτης Απόδοσης λόγω του μοτίβου που έχει, δηλαδή εάν έχουν πάρει δάνεια στο παρελθόν υπάρχει μεγάλη πιθανότητα να χρησιμοποιήσουν την συγκεκριμένη πλατφόρμα για την εύρεση καινούριο δανείου είτε για αυτούς είτε για άλλο μέλος της οικογένειάς τους.

- **Αναλογία επιπέδου εκπαίδευσης και της γνώση τους σε Η/Υ.**

Εδώ, ως ανεξάρτητη μεταβλητή, είναι μια μετρική (measure) που στοχεύει στην εύρεση της Αναλογίας μεταξύ του επιπέδου εκπαίδευσης των ερωτηθέντων σε σχέση με την γνώση τους σε Η/Υ. Αυτός ο Δείκτης Απόδοσης δημιουργήθηκε λόγω του μοτίβου που έχει, δηλαδή εάν κάποιος άνθρωπος έχει υψηλότερο επίπεδο εκπαίδευσης και γενικά γνώσεις σε Η/Υ, τότε έχει και μεγαλύτερη πιθανότητα να χρησιμοποιήσει το διαδίκτυο για την έρευνά του.

- **Σύνολο εργαζομένων ανά νοικοκυριό.**

Εδώ, ως ανεξάρτητη μεταβλητή, είναι μια μετρική (measure) που στοχεύει στην εύρεση του Συνόλου των ατόμων που εργάζονται σε ένα νοικοκυριό. Αυτός ο Δείκτης Απόδοσης δημιουργήθηκε λόγω του μοτίβου που έχει, δηλαδή εάν ένα νοικοκυριό έχει περισσότερα άτομα που εργάζονται έχει και μεγαλύτερη πιθανότητα να ξεπληρώσει κάποιο δάνειο και άρα και να παίρνει περισσότερα, οπότε και θα έχει μεγαλύτερη πιθανότητα να επισκεφθεί την πλατφόρμα που πρόκειται να δημιουργηθεί για την εύρεση ενός καινούριο δανείου.

3.5 Μετατροπή δεδομένων

3.5.1. Εξαρτημένη μεταβλητή – Κωδικοποίηση

Η εξαρτημένη μεταβλητή y , που αντιπροσωπεύει το Ποσοστό % των εν δυνάμει χρηστών της πλατφόρμας, για να εφαρμοστεί στο μοντέλο του Νευρωνικού Δικτύου με την βιβλιοθήκη 'sklearn' της γλώσσας Python, ήταν αναγκαίο να κωδικοποιηθεί. Δεδομένου ότι είναι μεταβλητή Τάξης, η κωδικοποίηση έγινε με τέτοιο τρόπο ώστε κάθε κατηγορία από τις πέντε (1: Ελάχιστα, 2: Λίγο, 3: Δεν ξέρω, 4: Πολύ, 5: Πάρα πολύ) να απέχει κατάλληλα σε σχέση με τις υπόλοιπες.

Ιδιαίτερα, η κωδικοποίηση αφορά μετατροπή της κάθε τιμής σε δυαδική τιμή (0,1) με πλήθος ψηφίων όσο και το πλήθος των κατηγοριών, δηλαδή πέντε δυαδικά ψηφία. Εδώ, για να ξεχωρίσει κάθε κατηγορία μεταξύ της, έχει σημασία η θέση του ψηφίου της μονάδας. Όσο πιο δεξιά είναι στην δυαδική τιμή εντοπίζεται η μονάδα, τόσο πιο μεγάλη είναι η τιμή στο δεκαδικό σύστημα αρίθμησης. Παράδειγμα το 1 είναι 10000, και το 5 είναι 00001¹.

3.5.2 Ανεξάρτητες μεταβλητές - Κανονικοποίηση

Οι ανεξάρτητες μεταβλητές X , που είναι πέντε στο πλήθος και αντιπροσωπεύουν κάποια χαρακτηριστικά του προβλήματος της έρευνας, για να εφαρμοστούν στο μοντέλο του Νευρωνικού Δικτύου με την βιβλιοθήκη 'sklearn' της γλώσσας Python, ήταν αναγκαίο να μετασχηματιστούν σε κανονικοποιημένη μορφή.

¹ <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.OneHotEncoder.html>

Πραγματοποιήθηκε δηλαδή μετασχηματισμός των δεδομένων για να κεντραριστούν γύρω από το μηδέν. Αυτό έγινε διότι από τα αρχικά δεδομένα μπορεί να μην παραχθούν τα κατάλληλα αποτελέσματα, διότι εάν ένα από τα χαρακτηριστικά έχει διακύμανση που είναι τάξης μεγέθους μεγαλύτερη από τα άλλα, μπορεί να κυριαρχήσει και έτσι μπορεί να μην μάθει ο εκτιμητής άλλα χαρακτηριστικά σωστά, όπως αναμένεται. Για τον λόγο αυτό τα δεδομένα μετασχηματίστηκαν να ακολουθούν την Κανονική κατανομή (Gaussian) με μηδενικό μέσο όρο και μοναδιαία διακύμανση. Η κανονικοποίηση πραγματοποιήθηκε αφαιρώντας την μέση τιμή κάθε χαρακτηριστικού και στην συνέχεια διαιρώντας τα μη σταθερά χαρακτηριστικά με την τυπική τους απόκλιση², με τον παρακάτω Τύπο 1:

$$z = (x - u) / s$$

Τύπος 1: Τύπος Κανονικοποίησης Ανεξάρτητων μεταβλητών

όπου:

x = δεδομένο εισόδους δείγματος

u = μέσος όρος δείγματος ή μηδέν

s = τυπική απόκλιση ή μονάδα

3.6 Διαχωρισμός δεδομένων σε δεδομένα train – test

Τα δεδομένα αφότου κωδικοποιήθηκαν και κανονικοποιήθηκαν αντίστοιχα, έγινε διαχωρισμός τους σε δεδομένα Εκμάθησης (Train data) και δεδομένα Ελέγχου (Test data). Η διαδικασία του διαχωρισμού πραγματοποιήθηκε με σκοπό την εκμάθηση των χαρακτηριστικών των δεδομένων από το μοντέλο (Model Training) και στην συνέχεια εφαρμόστηκε ο έλεγχος της αποτελεσματικότητας του μοντέλου με την χρήση των δεδομένων ελέγχου (Model Testing).

Ο διαχωρισμός των δεδομένων πραγματοποιήθηκε με την χρήση της βιβλιοθήκης ‘sklearn’ της γλώσσας Python³. Ιδιαίτερα, τα δεδομένα αυτά, διαχωρίστηκαν με τον εμπειρικό κανόνα εβδομήντα – τριάντα (Gholamy και άλλοι, 2018). Τα δεδομένα Εκμάθησης (Train data) αποτελούν το 70% του συνόλου, ενώ τα δεδομένα Ελέγχου (Test

² <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html>

³ https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.train_test_split.html

data) το 30%. Ο διαχωρισμός τους ήταν τυχαίος μεταξύ των δεδομένων και με σταθερό Seed.

3.7 Περιγραφή του μοντέλου Νευρωνικού δικτύου

Η παρούσα έρευνα ασχολείται με την συμβολή της μηχανικής μάθησης στο σχεδιασμό μιας στρατηγικής για βέλτιστη λήψη αποφάσεων και κατ' επέκταση του καθορισμού των ελάχιστων δυνατών λειτουργιών (features) που θα φέρει το προϊόν μιας ηλεκτρονικής πλατφόρμας από μια εταιρία τεχνολογίας. Για τον λόγο αυτό, επιλέχθηκε να δημιουργηθεί ένα μοντέλο μηχανικής μάθησης (Machine Learning) και συγκεκριμένα Βαθιάς Μάθησης (Deep Learning). Έτσι, στην έρευνα αυτή επιλέχθηκε να εφαρμοστεί ο αλγόριθμος του Νευρωνικού Δικτύου.

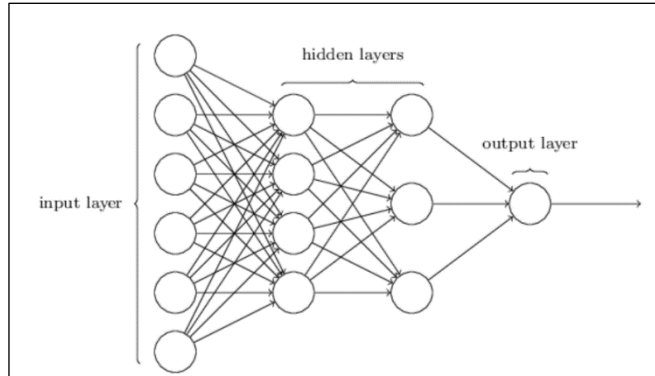
Το χαρακτηριστικό των τεχνητών Νευρωνικών Δικτύων είναι η αναγνώριση προτύπων με την μίμηση των Νευρωνικών Δικτύων του ανθρώπινου εγκεφάλου. Η βελτίωση της ταχύτητας των ηλεκτρονικών υπολογιστών έκανε πιο αποδοτική την προσομοίωση των νευρωνικών διαδικασιών (Fausett, 1993). Συγκεκριμένα, ένα Νευρωνικό Δίκτυο αποτελείται από ένα αφηρημένο μοντέλο διασυνδεδεμένων νευρώνων, των οποίων η ειδική διάταξή του και σύνδεσή του μπορεί να χρησιμοποιηθεί για την επίλυση υπολογιστικών προβλημάτων και να εφαρμοστεί σε διάφορους τομείς, όπως για παράδειγμα η οικονομία. Για να λυθούν αυτά τα προβλήματα, θα πρέπει τα Νευρωνικά Δίκτυα να εκπαιδεύονται (Kriesel, 2005).

Όσον αφορά την δομή και την λειτουργία τους, γενικά το αφηρημένο μοντέλο ενός Νευρωνικού Δικτύου αποτελείται από νευρώνες ή αλλιώς κόμβους. Οι κόμβοι, μπορούν να πάρουν πληροφορίες από το εξωτερικό του ή από άλλους κόμβους και να τις μεταδώσουν σε άλλους κόμβους και τέλος να τις εξάγουν ως τελικά αποτελέσματα. Έτσι, μπορεί να γίνει διάκριση μεταξύ τριών κόμβων, των κόμβων εισόδου (input nodes), των κρυμμένων κόμβων (hidden nodes) και των κόμβων εξόδου (output nodes).

Οι κόμβοι εισόδου λαμβάνουν πληροφορίες της μορφής ενός μοτίβου ή σήματος από τον εξωτερικό κόσμο. Οι κρυφοί κόμβοι βρίσκονται μεταξύ των κόμβων εισόδου και εξόδου και σχηματίζουν εσωτερικά μοτίβα πληροφορίας. Οι κόμβοι εξόδου αναμεταδίδουν την πληροφορία ή το σήμα στον εξωτερικό κόσμο ως αποτέλεσμα.

Επίσης, ένα ακόμα στοιχείο της δομής του Νευρωνικού Δικτύου είναι οι ακμές (edges), όπου συνδέουν τους κόμβους μεταξύ τους. Ανάλογα με την δύναμη και το νόημα της

σύνδεσης, η ακμή έχει μια συγκεκριμένη στάθμιση (weighting). Όσο μεγαλύτερη είναι η στάθμιση, τόσο μεγαλύτερη είναι και η επιρροή μπορεί να ασκήσει ένας κόμβος σε έναν άλλον συνδεδεμένο κόμβο. Η δομή ενός απλού Νευρωνικού Δικτύου απεικονίζεται στην *Εικόνα 4*:



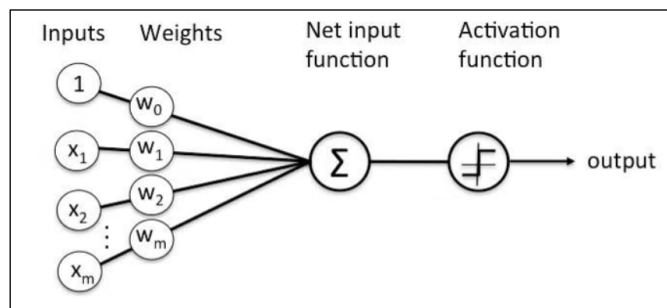
Εικόνα 4: Απλή δομή ενός Νευρωνικού Δικτύου

Ακόμη, η ροή της πληροφορίας σε ένα Νευρωνικό Δίκτυο μπορεί να πραγματοποιηθεί είτε από τους κόμβους εισόδου προς τους κόμβους εξόδου και ονομάζεται 'forward', είτε από τους κόμβους εισόδου προς τους κόμβους εξόδου και ονομάζεται 'backward'. Η πρώτη διαδικασία πραγματοποιείται στα Δίκτυα προώθησης (feed forward networks), ενώ η δεύτερη στα Επαναλαμβανόμενα δίκτυα (recurrent networks). Για να μπορέσει ένα Νευρωνικό Δίκτυο να χρησιμοποιηθεί, πρέπει πρώτα να εκπαιδευτεί. Έτσι, ανάλογα τα δεδομένα εκπαίδευσης και τους κανόνες μάθησης, το Νευρωνικό Δίκτυο βαραίνει τις συνδέσεις κατάλληλα μέχρι να αναπτύξει κάποια <<νοημοσύνη>> (Alsaadi και Mijwil, 2019).

Στην δεκαετία του 1950 με 1960, είχε αναπτυχθεί ένας τύπος τεχνητού νευρώνα που ονομάζεται 'Perceptron'. Ο τύπος αυτός όμως, εξαιτίας της δομής που παίρνει, παράγει μια binary τιμή για έξοδο και είναι αποτελεσματικός μόνο σε γραμμικά μοτίβα (Nielsen, 2019). Το γεγονός ότι στον πραγματικό κόσμο τα προβλήματα είναι συνήθως μη γραμμικά, έκανε αυτό το μοντέλο να μην είναι αποτελεσματικό σε αυτού του είδους τα προβλήματα. Έτσι, στην συνέχεια αναπτύχθηκαν νευρώνες που προσεγγίζουν καλύτερα τα πραγματικά μη γραμμικά προβλήματα βασισμένοι στην λογική του Perceptron.

Σημαντικό είναι όμως να αναφερθεί το πως δουλεύει ένα απλό μοντέλο Perceptron, δεδομένου ότι από αυτό δημιουργήθηκαν και άλλα είδη που χρησιμοποιούνται ευρέως σήμερα. Όσον αφορά την δομή του, το μοντέλο αυτό, ως είσοδο (input) μπορεί να λαμβάνει πολλές πραγματικές τιμές και να παράγει ως έξοδο (output) μία μόνο δυαδική (binary) τιμή. Κάθε μία από τις εισόδους (X_i) έχει ένα βάρος (w_i) που συσχετίζεται με αυτή. Τα

βάρη (weights), είναι αυτά που καθορίζουν την σημαντικότητα μιας εισόδου στην διαδικασία της λήψης αποφάσεων του μοντέλου. Ακόμη, προστίθεται σε αυτό η μεροληψία (bias). Το ποια θα είναι η δυαδική έξοδος του μοντέλου αποφασίζεται βάση ενός ορίου (W_0). Εάν το σταθμισμένο άθροισμα των εισόδων είναι μεγαλύτερο από το όριο (threshold), τότε η τιμή της εξόδου θα είναι '1', αλλιώς θα είναι '0'. Τέλος, για να είναι η έξοδος μέσα στο συγκεκριμένο εύρος $[0,1]$, χρησιμοποιείται η συνάρτηση ενεργοποίησης (activation function), μαζί με τις εισόδους, κατόπιν της μετατροπής τους σε σταθμισμένο άθροισμα σε μια συγκεκριμένη έξοδο που βασίζεται σε ένα σετ από κανόνες (Nielsen, 2019). Το μοντέλο Perceptron απεικονίζεται στην παρακάτω *Εικόνα 5* με τον αντίστοιχο τύπο του, *Τύπο 2*:



Εικόνα 5: Μοντέλο Perceptron

$$f(x) = \begin{cases} 1, & \text{εάν } w * x + b > 0 \\ 0, & \text{αλλιώς} \end{cases}$$

Τύπος 2: Τύπος εξαγωγής δεδομένων με το μοντέλο Perceptron με χρήση συνθήκης

όπου:

x = διάνυσμα δεδομένων εισόδου δείγματος (input)

w = διάνυσμα βαρών (weight)

b = μεροληψία (bias)

Επομένως, τα δεδομένα εξόδου προκύπτουν από τον *Τύπο 3*:

$$y = \sum_{i=1}^m (w_i * x_i) + bias$$

Τύπος 3: Τύπος εξαγωγής δεδομένων με το μοντέλο Perceptron

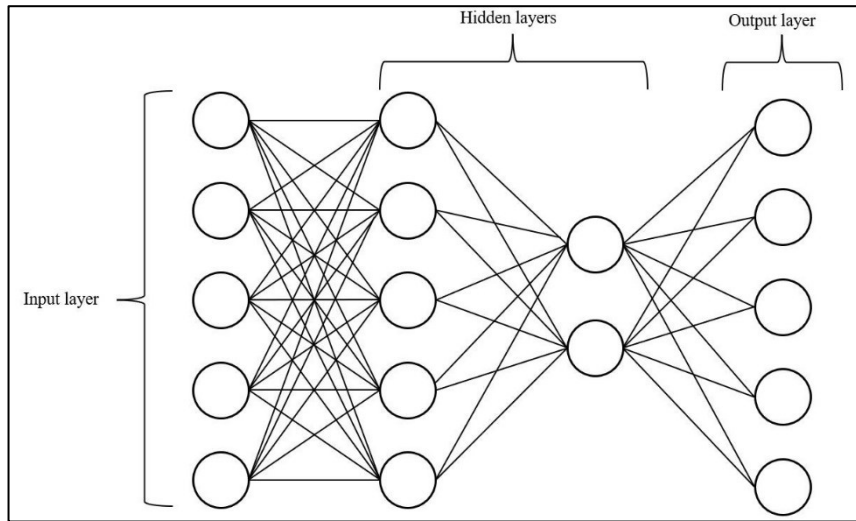
όπου:

y = διάνυσμα δεδομένων εξόδου δείγματος (output)

m = αριθμός των δεδομένων εισόδου του δείγματος

Σχετικά με την συνάρτηση ενεργοποίησης, υπάρχουν πολλά είδη όπως η Step function, η Softmax function, η Linear function, η Sigmoid function, η ReLu function κλπ. Η συνάρτηση ενεργοποίησης είναι αυτή που καθορίζει την τελική τιμή της εξόδου, διότι ένας νευρώνας δεν γνωρίζει το όριο (threshold) που τέθηκε και έτσι αποφασίζει εάν ένα συγκεκριμένος νευρώνας είναι ‘ενεργοποιημένος’ ή όχι (Szandała, 2021). Στην παρούσα έρευνα έγινε χρήση δύο συναρτήσεων ενεργοποίησης, της ReLu function και της Softmax function. Η επιλογή της πρώτης έγινε δεδομένου ότι αυτή η συνάρτηση επιτρέπει στα μοντέλα να μαθαίνουν πιο γρήγορα και να αποδίδουν καλύτερα και σύμφωνα με την βιβλιογραφία χρησιμοποιείται συνήθως ως συνάρτηση ενεργοποίησης για το πρώτο επίπεδο εισόδου (input layer) αλλά και για τα κρυφά επίπεδα (hidden layers) (Agarap, 2019).

Η επιλογή της δεύτερης έγινε διότι η συνάρτηση αυτή είναι κατάλληλη για κατηγορικές μεταβλητές και μάλιστα με πάνω από δύο κατηγορίες και σύμφωνα με την βιβλιογραφία χρησιμοποιείται συνήθως για το τελευταίο επίπεδο εξόδου (output layer) (Beale, 2014). Επίσης στην παρούσα έρευνα, δημιουργήθηκαν τέσσερα επίπεδα διότι σύμφωνα με την βιβλιογραφία, όσο λιγότερα δεδομένα έχει ένα πρόβλημα τόσο λιγότερα επίπεδα θα πρέπει να έχει, εδώ είναι 122 εγγραφές. Το πρώτο επίπεδο είναι το επίπεδο εισόδου (input layer), το δεύτερο και το τρίτο είναι τα δύο κρυφά επίπεδα (hidden layers) και το τέταρτο είναι το επίπεδο εξόδου (output layer). Στα πρώτα τρία επίπεδα, εισόδου και δύο κρυφά, χρησιμοποιήθηκε η ReLu function, ενώ στο επίπεδο εξόδου η Softmax function, για τους λόγους που αναφέρθηκαν παραπάνω. Ακόμη, στην έρευνα αυτή είναι πέντε οι κατηγορίες της εξόδου, είναι δηλαδή μια πολλαπλών τάξεων ταξινόμηση (multiclass classification). Στην παρακάτω Εικόνα 6, απεικονίζεται σε σχεδιάγραμμα η δομή των επιπέδων της παρούσας έρευνας.



Εικόνα 6: Σχεδιάγραμμα Δομής των Επιπέδων της έρευνας

Ιδιαίτερα, η ReLu function, αντιπροσωπεύει μια σχεδόν γραμμική συνάρτηση και επομένως διατηρεί τις ιδιότητες των γραμμικών μοντέλων, οι οποίες κάνουν το μοντέλο πιο εύκολο όσον αφορά την βελτιστοποίησή του με την μέθοδο Gradient-Descent. Επιπλέον, η ReLu function εκτελεί μια λειτουργία που βασίζεται σε (threshold) σε κάθε στοιχείο της εισόδου όπου οι τιμές είναι μικρότερες από το μηδέν τις μετατρέπει σε μηδέν (Szandala, 2021), εφαρμόζοντας τον Τύπο 4:

$$f(x) = \max(0, x) = \begin{cases} x_i, & \text{εάν } x_i \geq 0 \\ 0, & \text{εάν } x_i < 0 \end{cases}$$

Τύπος 4: Τύπος εξαγωγής δεδομένων από την ReLu function

όπου:

x_i = δεδομένο δείγματος

Από την άλλη, η Softmax function εφαρμόζεται για τον υπολογισμό της κατανομής πιθανοτήτων από ένα διάνυσμα πραγματικών αριθμών. Η Softmax function παίρνει είσοδο πολλαπλών τάξεων και παράγει μια έξοδο δυαδικών τιμών (binary) 0 ή 1. Επίσης, η συνάρτηση αυτή επιστρέφει πιθανότητες για κάθε κλάση, όπου η κλάση στόχος με την μεγαλύτερη πιθανότητα σε σχέση με τις υπόλοιπες, να είναι και αυτή που τελικά επικρατεί (Beale, 2014). Ο υπολογισμός της (Ijomah και άλλοι, 2018) γίνεται από τον Τύπο 5:

$$f(x_i) = \frac{\exp(x_i)}{\sum_j \exp(x_j)}$$

Τύπος 5: Τύπος εξαγωγής δεδομένων από την Softmax function

όπου:

x = δεδομένο δείγματος

Ένα ακόμα σημαντικό στοιχείο του Νευρωνικού μοντέλου είναι το είδος του. Το είδος του μοντέλου που επιλέχθηκε είναι το Διαδοχικό μοντέλο (Sequential model). Το Διαδοχικό μοντέλο εφαρμόστηκε στην παρούσα εργασία, ώστε να εισάγονται διαδοχικά τα δεδομένα στο μοντέλο του Νευρωνικού Δικτύου⁴. Αυτό σημαίνει ότι τα δεδομένα εξόδου από ένα επίπεδο εισάγονται ως δεδομένα εισόδου στο επόμενο επίπεδο. Σύμφωνα με την βιβλιογραφία αυτού του είδους το μοντέλο χρησιμοποιείται για προβλήματα ταξινόμησης (Zargar, 2021).

Άλλη μια συνάρτηση που χρησιμοποιήθηκε στο παρόν μοντέλο του Νευρωνικού Δικτύου είναι η Συνάρτηση Απώλειας (Loss function), με σκοπό την μέτρηση της απόδοσής του. Δεδομένου ότι η εκπαίδευση ενός Νευρωνικού Δικτύου είναι ένα είδος προβλήματος βελτιστοποίησης, η Συνάρτηση Απώλειας, θα πρέπει να ελαχιστοποιηθεί. Γενικά, οι Συναρτήσεις Απώλειας καθορίζονται βάση το τι χρειάζεται να μάθει το μοντέλο σύμφωνα με κάποια κριτήρια (Bircanoğlu, 2017). Έτσι, το σφάλμα (loss) της τρέχουσας κατάστασης θα πρέπει να εκτιμάται επανειλημμένα, ώστε τα βάρη (weights) να μπορούν να ενημερώνονται για να μειωθεί η απώλεια της επόμενης αξιολόγησης. Στην έρευνα αυτή, η Συνάρτηση Απώλειας που χρησιμοποιήθηκε είναι η Categorical Cross Entropy function. Η Categorical Cross Entropy function εφαρμόστηκε διότι σύμφωνα με την βιβλιογραφία εφαρμόζεται σε προβλήματα ταξινόμησης πολλαπλών τάξεων. Πιο συγκεκριμένα, έχει σκοπό να φέρει τα στοιχεία της ίδιας κλάσης σε έναν N – διάστατο χώρο με συγκεκριμένα όρια (regions). Η εξαρτημένη μεταβλητή y, θα πρέπει να είναι κωδικοποιημένη, όπως εξηγήθηκε στο κεφάλαιο 3.5.1, ώστε να μπορεί να χρησιμοποιηθεί στην Categorical Cross Entropy function. Ιδιαίτερα, υπολογίζει την μέση διαφορά μεταξύ της πραγματικής (y_i) και της προβλεπόμενης (ŷ_i) κατανομής πιθανοτήτων για όλες τις κατηγορίες του προβλήματος και έτσι η βαθμολογία της (score) ελαχιστοποιείται. Όσο η

⁴ [The Sequential model \(keras.io\)](#)

τιμή της πλησιάζει στο μηδέν, τόσο καλύτερη είναι η ελαχιστοποίηση (Gentile, 1998). Ο υπολογισμός του πραγματοποιείται μέσα από τον παρακάτω Τύπο 6:

$$\text{Cross Entropy} = - \frac{1}{N} \sum_i y_i * \log(\hat{y}_i)$$

Τύπος 6: Τύπος μέτρησης της Απόδοσης του μοντέλου με την Categorical Cross Entropy function

όπου:

N = πλήθος δεδομένων δείγματος

y_i = πραγματική κατηγορία

\hat{y}_i = προβλεπόμενη κατηγορία

Για να μπορέσει να μετρηθεί η απόδοση του Νευρωνικού Δικτύου μέσω της Συνάρτησης απόδοσης, εφαρμόζεται ένας αλγόριθμος βελτιστοποίησης (Optimizer). Ένας βελτιστοποιητής είναι ένας αλγόριθμος ή μια μέθοδος που χρησιμοποιείται για να αλλάξει τα χαρακτηριστικά του Νευρωνικού Δικτύου, όπως για παράδειγμα τα βάρη (weights) και τον ρυθμό εκμάθησης, με σκοπό να μειωθούν οι απώλειες (losses). Η λογική του για να βρει το ελάχιστο είναι να <<κυλίσει>> με έναν πιο αργό τρόπο ώστε να κάνει μια προσεχτική αναζήτηση και να μην προσπεράσει το ελάχιστο (Keren, 2019). Στην περίπτωση αυτή χρησιμοποιήθηκε ο βελτιστοποιητής 'Adam'. Υπολογίζεται με τον παρακάτω Τύπος 7 – 8:

$$m_t = \frac{m_t}{1 - \beta_1^t}$$

Τύπος 7: Τύπος βελτιστοποιητή Adam της πρώτης στιγμής

$$u_t = \frac{u_t}{1 - \beta_2^t}$$

Τύπος 8: Τύπος βελτιστοποιητή Adam της δεύτερης στιγμής

όπου:

m_i = αρχικός μέσος όρος

\hat{m}_i = μέσος όρος πρώτης στιγμής

u_i = αρχική διακύμανση αβεβαιότητας

\hat{u}_i = διακύμανση αβεβαιότητας δεύτερης στιγμής

Ακόμη μια παράμετρος που εφαρμόστηκε είναι τα epochs, δηλαδή ο αριθμός των επαναλήψεων της εκπαίδευσης του μοντέλου του Νευρωνικού Δικτύου. Τα epochs καθορίζουν το πλήθος της ενημέρωσης των βαρών (weights) σε ένα Νευρωνικό Δίκτυο. Ουσιαστικά, σε κάθε επανάληψη εκπαίδευσης του μοντέλου ο βελτιστοποιητής (Optimizer) αναζητά το ελάχιστο σημείο, διότι ένας βελτιστοποιητής είναι επαναληπτικός αλγόριθμος όπου αναζητά επαναληπτικά το ελάχιστο σημείο. Σύμφωνα με την βιβλιογραφία, δεν υπάρχει κάποιος σταθερός αριθμός του πλήθους των epochs που θα δώσει το βέλτιστο μοντέλο, και έτσι για να βρεθεί το βέλτιστο πλήθος θα πρέπει να γίνουν δοκιμές στο Νευρωνικό Δίκτυο (Afaq, 2020). Στην παρούσα εργασία, μετά από διάφορες δοκιμές εφαρμόστηκαν 1000 epochs, δηλαδή επαναλήψεις για την αναζήτηση του ελάχιστου σημείου από τον βελτιστοποιητή. Η παράμετρος αυτή, μαζί με τα δεδομένα Εκπαίδευσης (Train data) και δεδομένα Ελέγχου (Test data) χρησιμοποιήθηκαν κατά την προσαρμογή του μοντέλου του Νευρωνικού Δικτύου (Model fitting).

3.8 Confusion matrix

Για την αξιολόγηση του μοντέλου και ιδιαίτερα για εύρεση των σωστών και λάθος ταξινομήσεων των κατηγοριών της εξαρτημένης μεταβλητής y , έγινε χρήση του Πίνακα Σύγχυσης (Confusion Matrix). Ο Πίνακας Σύγχυσης, είναι ένα εργαλείο για την επαλήθευση των αποτελεσμάτων της ταξινόμησης, δηλαδή την εξακρίβωση της συνολικής ακρίβειας της ταξινόμησης, αλλά και την εύρεση των σωστών και μη ταξινομήσεων της μεταβλητής y .

Στον Πίνακα αυτό, σε ορθά ταξινομημένα στοιχεία, τα οποία βρίσκονται διαγωνίως του πίνακα, παράγεται ένας κοινός μέσος όρος επί τις εκατό και έτσι προκύπτει η συνολική ακρίβεια της ταξινόμησης. Επίσης, μπορεί να εξεταστεί το ποσοστό συνολικής ακρίβειας ταξινόμησης ανά κατηγορία (Ταγκούτα, 2017). Στην οριζόντια πλευρά (στήλες) περιέχει τις Προβλεπόμενες κατηγορίες, ενώ στην κάθετη πλευρά περιέχει τις Πραγματικές κατηγορίες (Doulamis και άλλοι, 2021). Στην *Εικόνα 7*, απεικονίζεται ένας Πίνακας Σύγχυσης και στον *Τύπο 9*, ο υπολογισμός της Απόδοσης (Accuracy) του μοντέλου Νευρωνικού Δικτύου μέσα από αυτό.

		Προβλεπόμενες Κατηγορίες			
		<i>Class₁</i>	<i>Class₂</i>	...	<i>Class_m</i>
Πραγματικές Κατηγορίες	<i>Class₁</i>	n_{11}	n_{12}	...	n_{1m}
	<i>Class₂</i>	n_{21}	n_{22}	...	n_{2m}

	<i>Class_m</i>	n_{m1}	n_{m2}	...	n_{mm}

Εικόνα 7: Παράδειγμα ενός γενικού Πίνακα Σύγκρισης (Confusion Matrix)

$$Accuracy = \frac{n_{11} + n_{22} + \dots + n_{mm}}{N} * 100$$

Τύπος 9: Τύπος υπολογισμού της Απόδοσης (Accuracy) του μοντέλου

όπου:

n = δεδομένο

N = πλήθος δεδομένων

m = πλήθος κλάσεων

4. ΑΠΟΤΕΛΕΣΜΑΤΑ

Κατά την χρήση του μοντέλου Νευρωνικού Δικτύου και την επιλογή των λειτουργιών/χαρακτηριστικών (features) από τους ερωτηθέντες, προέκυψαν τα παρακάτω αποτελέσματα. Όσον αφορά το κυρίως κομμάτι της έρευνας, δηλαδή της κατασκευής του Νευρωνικού Δικτύου, όπως παρατηρείται και στην Εικόνα 8, η Απόδοσή του προέκυψε είναι αρκετά μεγάλη της τάξης του 91,89%.

$$NN Accuracy = 0.918918918918919$$

Εικόνα 8: Απόδοσης (Accuracy) του μοντέλου Νευρωνικού Δικτύου

Ακόμη, στο Διάγραμμα της *Εικόνας 9*, παρατηρείται ότι η πλειοψηφία των ερωτηθέντων (48/122) απάντησε ότι θα χρησιμοποιούσε μια πλατφόρμα χρηματοοικονομικών υπηρεσιών, σε βαθμό 5: Πάρα Πολύ (με εύρος 1: Ελάχιστα – 5: Πάρα πολύ) . Υπάρχει όμως και ένα μικρό μέρος του πληθυσμού που θα το χρησιμοποιούσε 1: Ελάχιστα (18/122). Επίσης, από τον μέσο όρο του (3.8) προκύπτει ότι οι περισσότεροι ερωτηθέντες θα χρησιμοποιούσαν της πλατφόρμα και συγκεκριμένα 7/10 ανθρώπους θα την χρησιμοποιούσαν.



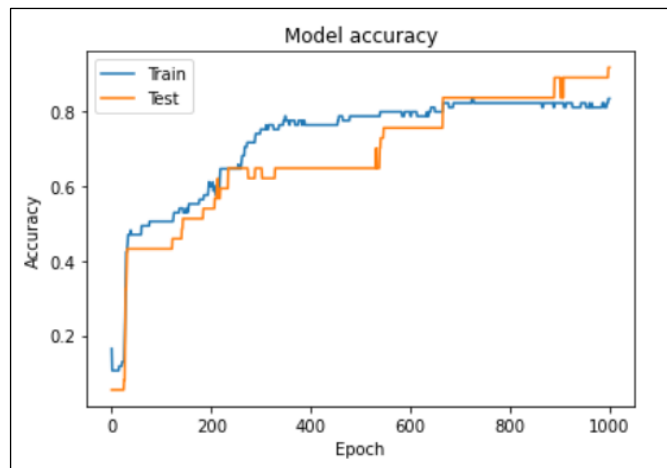
Εικόνα 9: Ραβδόγραμμα με την Εξαρτημένη μεταβλητή y

Σχετικά με τον Πίνακα Σύγκρισης, του *Πίνακα 3*, παρατηρείται ότι η Συνολική Απόδοση (Accuracy) είναι 91,89%. Επίσης, παρατηρείται ότι σε όλες τις κατηγορίες (1: Ελάχιστα, 2: Λίγο, 3: Δεν ξέρω, 4: Πολύ, 5: Πάρα πολύ) τα δεδομένα ταξινομήθηκαν με πολύ υψηλό ποσοστό επιτυχίας. Ακόμη, στην κλάση 1 τα δεδομένα φαίνεται να έχουν τέλεια ταξινόμηση με ποσοστά 100%. Επίσης στην κατηγορία 4 υπάρχουν οι περισσότερες λάθος ταξινομήσεις με ποσοστό ταξινόμησης 80%, όπου το μοντέλο συγχέει την κατηγορία αυτή με την 2, 3. Αλλά και την κατηγορία 5 το μοντέλο την συγχέει με την κατηγορία 4, με ποσοστό σωστής ταξινόμησης 93.33%. Οι κατηγορίες 2 και 3 παρουσιάζουν σωστή ταξινόμηση όσον αφορά τα πραγματικά δεδομένα τους, όμως κάποια δεδομένα που τους έχουν ταξινομηθεί σε αυτές τις τάξεις στην πραγματικότητα ανήκουν σε άλλες, όπως την προηγούμενη περίπτωση της κατηγορίας 5.

Πίνακας 3: Πίνακας Σύγκρισης του μοντέλου του Νευρωνικού Δικτύου

Προβλεπόμενες Κατηγορίες	Προβλεπόμενες Κατηγορίες					Ποσοστό Σωστών	
	1	2	3	4	5		
Πραγματικές Κατηγορίες	1	6	0	0	0	0	100%
	2	0	2	0	0	0	100%
	3	0	0	4	0	0	100%
	4	0	1	1	8	0	80%
	5	0	0	0	1	14	93.33%
Ποσοστό Σωστών		100%	66.66%	80%	88.88%	100%	91,89%

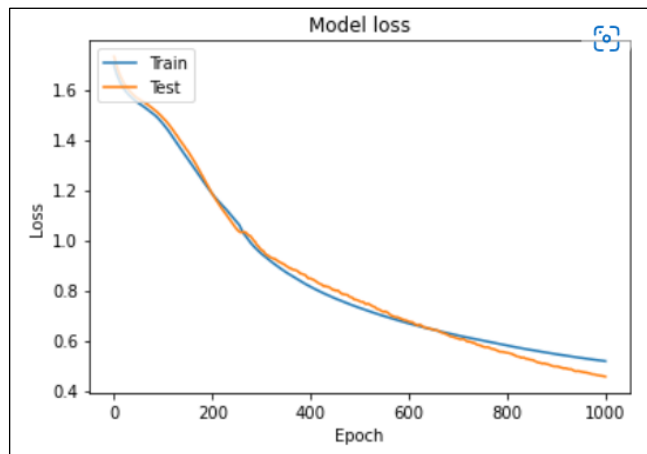
Από πλευράς του μοντέλου του Νευρωνικού Δικτύου, στην *Εικόνα 10*, παρατηρείται ότι όσο αυξάνονται ο αριθμός των επαναλήψεων (epochs) της εκπαίδευσης του μοντέλου του Νευρωνικού Δικτύου, τόσο αυξάνεται και το ποσοστό της απόδοσης του μοντέλου. Ακόμη, φαίνεται ότι από ένα σημείο και μετά το ποσοστό της απόδοσης τείνει να σταθεροποιείται γύρω στο 80%.



Εικόνα 10: Ποσοστό Απόδοσης (Accuracy) του μοντέλου Νευρωνικού Δικτύου σε σχέση με τον αριθμό των επαναλήψεων της εκπαίδευσης (epochs)

Επιπροσθέτως, όσον αφορά την Συνάρτηση Απώλειας (Loss function) που στην παρούσα έρευνα είναι η Categorical Cross Entropy function, στην *Εικόνα 11* παρατηρείται ότι όσο αυξάνεται ο αριθμός των επαναλήψεων (epochs) της εκπαίδευσης του μοντέλου

του Νευρωνικού Δικτύου, τόσο ελαχιστοποιείται και η Συνάρτηση Απώλειας (Loss function).



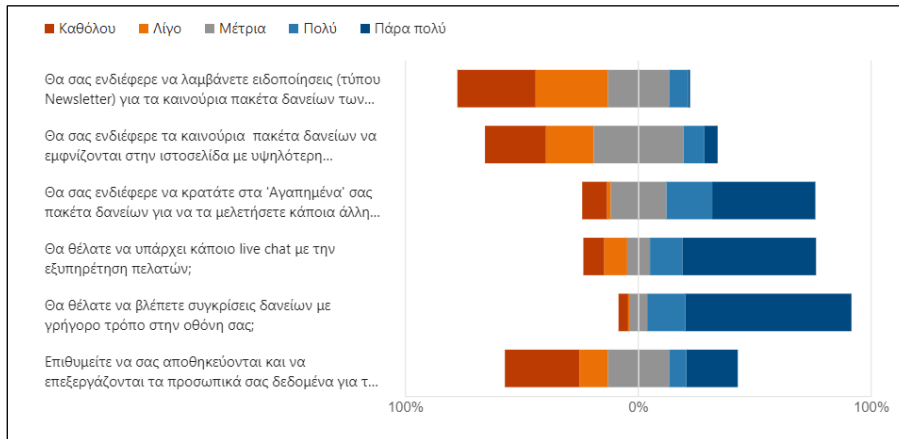
Εικόνα 11: Συνάρτηση Απώλειας (Loss function) του μοντέλου Νευρωνικού Δικτύου σε σχέση με τον αριθμό των επαναλήψεων της εκπαίδευσης (epochs)

Τέλος, σχετικά με τα χαρακτηριστικά / λειτουργίες (features) της ηλεκτρονικής πλατφόρμας χρηματοοικονομικών υπηρεσιών που πρόκειται να περιέχει όταν δημιουργηθεί, με σειρά προτεραιότητας. Οι ερωτηθέντες απάντησαν στην ερώτηση <<Δηλώστε την σημαντικότητα που έχουν για εσάς τα παρακάτω χαρακτηριστικά των ιστοσελίδων χρηματοοικονομικών υπηρεσιών>> [1: Καθόλου, 2: Λίγο, 3: Μέτρια, 4: Πολύ, 5: Πάρα πολύ]. Μεταξύ των έξι χαρακτηριστικών / λειτουργιών τρία από αυτά ήταν τα πιο δημοφιλή, όπως φαίνεται και στην *Εικόνα 12*. Τα τρία δημοφιλέστερα χαρακτηριστικά/λειτουργίες είναι:

1. Θα θέλατε να βλέπετε συγκρίσεις δανείων με γρήγορο τρόπο στην οθόνη σας;
2. Θα σας ενδιέφερε να κρατάτε στα 'Αγαπημένα' σας πακέτα δανείων για να τα μελετήσετε κάποια άλλη στιγμή;
3. Θα θέλατε να υπάρχει κάποιο live chat με την εξυπηρέτηση πελατών;

Ενώ τα τρία λιγότερο δημοφιλή χαρακτηριστικά / λειτουργίες προέκυψε ότι είναι:

1. Θα σας ενδιέφερε να λαμβάνετε ειδοποιήσεις (τύπου Newsletter) για τα καινούρια πακέτα δανείων των τραπεζών;
2. Θα σας ενδιέφερε τα καινούρια πακέτα δανείων να εμφανίζονται στην ιστοσελίδα με υψηλότερη προτεραιότητα από τα παλαιότερα;
3. Επιθυμείτε να σας αποθηκεύονται και να επεξεργάζονται τα προσωπικά σας δεδομένα για την πρόταση εξατομικευμένων πακέτων δανείων;



Εικόνα 12: Χαρακτηριστικά/λειτουργίες (features) της ηλεκτρονικής πλατφόρμας

5. ΣΥΖΗΤΗΣΗ

Σε αυτή την ερευνητική εργασία, πραγματοποιήθηκε η εκτίμηση της επιτυχίας ενός online προϊόντος με Δείκτες Απόδοσης (KPIs) εφαρμόζοντας μεθόδους μηχανικής μάθησης. Από την έρευνα εντοπίστηκε ότι σύμφωνα με τις απαντήσεις που είχαν δοθεί στο ερωτηματολόγιο, ο αλγόριθμος παρουσίασε υψηλό ποσοστό εγκυρότητας σχετικά με το αν το μεγαλύτερο μέρος του πληθυσμού που ενδιαφέρεται για απόκτηση δανείου, θα χρησιμοποιούσε την εν λόγω πλατφόρμα για την εύρεση του καταλληλότερου δανείου.

Η σημαντικότητα αυτής της έρευνας έγκειται στην ελαχιστοποίηση του οικονομικού ρίσκου που μπορεί να σημειωθεί μέσα από την συγκεκριμένη προσέγγιση που προβάλλεται σε αυτή την εργασία η οποία εστιάζει στην εκτίμηση της επιτυχίας της πλατφόρμας πριν αυτή υλοποιηθεί. Με αυτό τον τρόπο παρέχεται η δυνατότητα να επιτευχθεί ευστοχότερος σχεδιασμός που θα έχει ως αποτέλεσμα την αύξηση της πιθανότητας να πετύχει στην αγορά μετά την υλοποίησή της. Η εστίαση της εργασίας στο στάδιο πριν την υλοποίηση της πλατφόρμας είναι ένα χαρακτηριστικό που τη

διαφοροποιεί από άλλες παρόμοιες ερευνητικές προσπάθειες οι οποίες χρησιμοποιούσαν μηχανική μάθηση με Δείκτες Απόδοσης μετά την δημιουργία ενός προϊόντος.

Τα συμπεράσματα της μεθόδου που χρησιμοποιήθηκε καθώς και ο τρόπος που συνεισέφερε η μηχανική μάθηση στη λήψη αποφάσεων και απόκτηση γνώσης από τα δεδομένα του ερωτηματολογίου μπορούν να φανούν χρήσιμα σε περιπτώσεις που μια επιχείρηση πραγματοποιεί ελέγχους στην αγορά για να εντοπίσει κατά πόσο μια ηλεκτρονική πλατφόρμα που ετοιμάζεται, έχει τα απαραίτητα χαρακτηριστικά για γίνει ανταγωνιστική. Επίσης, προσφέροντας μια παραπάνω σιγουριά ότι το συγκεκριμένο ηλεκτρονικό προϊόν πρόκειται να προσελκύσει το ενδιαφέρον των καταναλωτών έχει ως αποτέλεσμα να διευκολύνει την συγκέντρωση χρηματοδότησης από επενδυτές επειδή θα έχουν ενδείξεις ότι το προϊόν στο οποίο επενδύουν εξυπηρετεί τις ανάγκες μεγάλου μέρους του πληθυσμού. Η συγκέντρωση χρηματοδότησης είναι υψίστης σημασίας στο ξεκίνημα ενός νέου προϊόντος και σε αντίθεση με παρόμοιες ερευνητικές εργασίες που χρησιμοποιούν μηχανική μάθηση με Δείκτες Απόδοσης μετά την δημιουργία ενός προϊόντος, στη συγκεκριμένη εργασία πραγματοποιείται πριν την υλοποίηση για να επιτευχθεί καλύτερος σχεδιασμός και λιγότερο ρίσκο.

Μελλοντικές προεκτάσεις έρευνας που προκύπτουν από αυτή την εργασία όσον αφορά το μοντέλο Νευρωνικού Δικτύου είναι η εφαρμογή της μεθόδου Διασταυρούμενης Επικύρωσης (cross validation) για την εκπαίδευση πολλών μοντέλων σε υποσύνολα των διαθέσιμων δεδομένων εισόδου και την αξιολόγησή τους στο συμπληρωματικό υποσύνολο των δεδομένων με σκοπό τον καλύτερο έλεγχο της ικανότητας του ταξινομητή να γενικεύει. Επιπλέον, όσον αφορά το επιχειρησιακό πλάνο, μπορεί να διερευνηθεί περισσότερο πως μπορούν να εφαρμοστούν τα ευρήματα από την ανάλυση με την μηχανική μάθηση στο να τεθούν στόχοι και να γίνουν επίσης σημαντικές αλλαγές στο επιχειρησιακό μοντέλο.

Συνεπώς, η χρήση της μηχανικής μάθησης στις επιχειρήσεις έχει αποδείξει ότι έχει την δυνατότητα να βελτιώσει πολλές διαδικασίες προσφέροντας ταχύτητα, ακρίβεια και δημιουργικότητα, ανακαλύπτοντας μοτίβα στα δεδομένα που χρησιμοποιήθηκαν κατά την ανάλυση. Ο συνδυασμός αυτών των ικανοτήτων είναι ιδιαίτερα χρήσιμος για την λήψη αποφάσεων σε μια επιχείρηση όπου πρέπει να λαμβάνονται υπόψιν πολλές παράμετροι για να υπάρξει μια απόφαση που είναι εξατομικευμένη και θα οδηγήσει στην βέλτιστη λύση για ικανοποίηση των πελατών και κατ' επέκταση αύξηση των κερδών.

6. ΣΥΜΠΕΡΑΣΜΑΤΑ

Τα μοντέλα Μηχανικής Μάθησης, σύμφωνα με την βιβλιογραφία, εφαρμόζονται για την αυτοματοποίηση της ανάλυσης των δεδομένων, την εύρεση των μοτίβων τους, την εκτίμηση τους, με τελικό σκοπό την λήψη αποφάσεων πάνω σε ένα πρόβλημα. Όσον αφορά τις επιχειρήσεις, εφαρμόζοντας τα παραπάνω, βοηθάει στην βελτίωση της παραγωγής και την οικονομικής τους ανάπτυξη μέσα από την δημιουργία καλύτερης στρατηγικής. Επίσης, μέσω της αυτοματοποίησης, αφήνει περισσότερο χρόνο στους αναλυτές να ασχοληθούν με άλλα ζητήματα σχετικά με το επιχειρησιακό πλάνο και έτσι γίνεται ακόμα καλύτερη λήψη αποφάσεων σχετικά με την ανάπτυξη της επιχείρησης. Ακόμη, μειώνει το κόστος και βελτιώνει την εμπειρία του πελάτη.

Ιδιαίτερα τα μοντέλα Νευρωνικών Δικτύων, σύμφωνα με την βιβλιογραφία, είναι εύκολα στην χρήση και είναι ιδανικά για προβλήματα ταξινόμησης. Εξαιτίας της φύσης τους, που παρομοιάζουν τα βιολογικά Νευρωνικά Δίκτυα, με τα πολλαπλά επίπεδά τους είναι ικανά να λύνουν πιο εύκολα τα προβλήματα τα οποία λύνονται πολύ δύσκολα με τις παραδοσιακούς μεθόδους ανάλυσης. Το γεγονός αυτό τα καθιστά πολύ δημοφιλή στις επιχειρήσεις.

Ο σκοπός της παρούσας έρευνας είναι να εκτιμήσει την επιτυχία της πλατφόρμας, χρησιμοποιώντας τις απαντήσεις των ερωτηθέντων και φτιάχνοντας στην συνέχεια ένα μοντέλο μηχανικής μάθησης. Για την εκτίμηση της επιτυχίας της πλατφόρμας χρησιμοποιήθηκαν Δείκτες Απόδοσης (Key Performance Indicators – KPIs) που βασίζονται στα δεδομένα που συλλέχθηκαν από τα ερωτηματολόγια, τα οποία στην συνέχεια εκτιμήθηκαν με την χρήση της μηχανικής μάθησης. Έτσι, δημιουργήθηκαν έξι KPIs από τα δεδομένα που συλλέχθηκαν και εφαρμόστηκε ένα μοντέλο Νευρωνικού Δικτύου όπου έγινε ταξινόμηση των κατηγοριών της εξαρτημένης μεταβλητής y (KPI1) με την χρήση των ανεξάρτητων μεταβλητών X_i (KPI2 – KPI6). Επίσης, από τα ερωτηματολόγια προέκυψε ότι 7/10 ανθρώπους θα χρησιμοποιούσαν αυτή την ηλεκτρονική πλατφόρμα, δηλαδή θα έχει επιτυχία η πλατφόρμα. Με την χρήση όμως της μηχανικής μάθησης, προέκυψε ότι κατά 91,89% είναι σίγουρο ότι θα έχει επιτυχία το προϊόν της επιχείρησης. Επίσης, παρατηρήθηκε ότι η κατηγορία που ο αλγόριθμος συγγέεται περισσότερο σε σχέση με τις υπόλοιπες είναι η κατηγορία 4 (4: Πολύ), πιθανότατα αξαιτίας των κοινών χαρακτηριστικών αυτής της κατηγορίας σε σχέση με τις υπόλοιπες.

Επιπλέον, όσον αφορά τα χαρακτηριστικά/λειτουργίες της πλατφόρμας, τρία ήταν τα πιο δημοφιλή χαρακτηριστικά από τα συνολικά έξι. Έτσι, αυτά τα τρία πρώτα είναι που

θα μπουν σε προτεραιότητα για να δημιουργηθούν σε αυτό το προϊόν, ώστε να προσελκύσει τους υποψήφιους χρήστες από την αρχή της δημιουργίας του.

Συνοψίζοντας, και βάση όλων των αποτελεσμάτων της παρούσας έρευνας η χρήση της μηχανικής μάθησης στις επιχειρήσεις είναι πολύ σημαντική. Θα πρέπει να γίνεται με συνδυασμό των σωστών μεθόδων διαχείρισης των δεδομένων που πραγματοποιείται πριν την δημιουργία μοντέλων, την σωστή ερμηνεία των αποτελεσμάτων κατόπιν την δημιουργίας τους, από τους κατάλληλους αναλυτές με σκοπό την ευημερία της επιχείρησης. Τέλος, είναι σημαντικό μια επιχείρηση να ακολουθεί την τεχνολογική πρόοδο ώστε συνεχώς να αποκομίζει τα οφέλη της και να αποκτά ανταγωνιστικό πλεονέκτημα.

7. ΒΙΒΛΙΟΓΡΑΦΙΑ

- Afaq, S. (2020). Significance Of Epochs On Training A Neural Network. *International Journal of Scientific & Technology Research*.
- Agarap, M. (2019). Deep Learning using Rectified Linear Units. *Arxiv*.
- Awan, H. (2014). The key performance indicators (KPIs) and their impact on overall organizational performance. *Quality & Quantity: International Journal of Methodology, Springer, vol. 48(6)*, 3127-3143.
- Bakker, E., Zheng, J., & Knight, L. (2008). Putting e-commerce adoption in a supply chain context. *International Journal of Operations & Production Management 28(4)*, 313–330.
- Ballog, K. (2020). The concept and competitiveness of agile organization in the fourth industrial revolution's drift. *International Scientific Conference Strategic Management and Decision Support Systems in Strategic Management*.
- Borges, M., Hoppen, N., & Luce, F. (2009). Information technology impact on market orientation in e-business. *Journal of Business Research 62(9)*, 883–890.
- Božič, ., K., & Dimovski, V. (2019). Business intelligence and analytics use, innovation ambidexterity, and firm performance: A dynamic capabilities perspective. *The Journal of Strategic Information Systems*, 28.
- Bryman, A. (2004). *Social Research Methods*. Oxford University Press.
- Brzozowska, A. (2015). E-business as a new trend in the economy. *International Conference on Communication, Management and Information Technology (ICCMIT)* (σσ. 1097-1098). Elsevier.
- Burney, D. S. (2008). Inductive & deductive research approach. Karachi, Pakistan.
- Chollet, F. (2018). *Deep Learning with Python*. Manning.
- Chollet, F. (2020). *The Sequential model*. Ανάκτηση από keras.io:
https://keras.io/guides/sequential_model/
- Dörnyei, Z. (2007). *Research Methods in Applied Linguistics*. Oxford University Press.
- Duan, Y. (2019). Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda. *International Journal of Information Management*.
- Ekinci, Y. (2015). *Designing Research Questionnaires for Business and Management Students*. Sage.
- Fausett, L. V. (1994). *Fundamentals of Neural Networks: Architectures, Algorithms And Applications*. Pearson.
- Fawcett, T., & Provost, F. (2013). *Data Science for Business*. O'Reilly Media, Inc.
- Gholamy, A., Kreinovich, V., & Kosheleva, O. (2018). Why 70/30 or 80/20 Relation Between Training and Testing Sets: A Pedagogical Explanation. *Departmental Technical Reports Computer Science at ScholarWorks@UTEP*.
- Goodfellow, I. (2016). *Deep Learning*. MIT Press.
- IBM. (2020). *Supervised Learning*. Ανάκτηση από IBM:
<https://www.ibm.com/cloud/learn/supervised-learning>
- Kitsios, F., & Kamariotou, M. (2019). Digital Business Strategy and Information Systems Planning: Determinants of Success.
- Kitsios, F. (2021). Artificial Intelligence and Business Strategy towards Digital Transformation: A Research Agenda. *Sustainability*, 1.
- Kriesel, D. (2005). A brief Introduction to Neural Networks. Bonn, Germany.
- Magumba, M. A. (2016). WEB 2.0 For Small And Medium Sized Companies:. 12.
- Markoulidakis, I., Rallis, I., Georgoulas, I., Kopsiaftis, G., Doulamis, A., & Doulamis, N. (2021). Multiclass Confusion Matrix Reduction Method and Its Application on Net Promoter Score Classification Problem. *The 14th Pervasive Technologies Related to Assistive Environments Conference* (σσ. 412–419). Corfu, Greece: Association for Computing Machinery.
- Mijwel, M. (2019, April). *Overview of Neural Networks*. Ανάκτηση από www.researchgate.net:
https://www.researchgate.net/profile/Maad-Mijwil/publication/332655457_Overview_of_Neural_Networks/links/5cc1f626299bf120977f7989/Overview-of-Neural-Networks.pdf

- Mishra , P., Pandey , C. M., & Singh, U. (2018). Scales of Measurement and Presentation of Statistical Data, . *Annals of Cardiac Anaesthesia*, 419–422.
- Norhayate, W., & Rahim, M. A. (2015). The Case Study Method in Business. *Sch. J. Arts. Humanit. Soc. Sci*, 105-109 .
- Nwankpa, C. E., Ijomah, W., Gachagan, A., & Marshall, S. (2018). Activation Functions: Comparison of Trends in Practice and Research for Deep Learning. *2nd International Conference on Computational Sciences and Technology* (σσ. 124 - 133). Arxiv.
- Pedregosa, F. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 2825-2830. Ανάκτηση από scikit-learn.org: <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.OneHotEncoder.html>
- Sauce, B., & Matzel, L. D. (2017). Inductive Reasoning. Στο *Encyclopedia of Animal Cognition and Behavior* (σσ. 1-8). Springer.
- Scikit-learn developers. (2021). Ανάκτηση από Standard Scaler: <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html>
- Scikit-learn developers. (2021). *One Hot Encoder*. Ανάκτηση από scikit-learn.org: <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.OneHotEncoder.html>
- Sujata, J., & Aniket, D. (2019). Artificial Intelligence Tools for Enhancing Customer Experience.
- Szandała, T. (2021). Review and Comparison of Commonly Used Activation Functions for Deep Neural Networks. *Arxiv*.
- Villazón, C. C., Pinilla, L. S., Olaso, J. R., Gandarias, N. T., & de Lacalle, N. L. (2020, July 24). Identification of Key Performance Indicators in Project-Based Organisations through the Lean Approach . *Sustainability*, σσ. 12(15), 5977.
- Whig, P. (2019). Artificial Intelligence and Machine Learning In Business. *International Journal on Integrated Education*.
- Whitney, H. (2012). *Data Insights*.
- Young, T. J. (2016). Questionnaires and Surveys. Στο Z. Hua, *Research Methods in Intercultural Communication: A Practical Guide* (σσ. 163-180). Wiley.
- Zargar, S. A. (2021). *Introduction to Sequence Learning Models: RNN, LSTM, GRU*.
- Ταγκούτα, Ε. (2017). *Αντικειμενοστραφής ταξινόμηση δορυφορικών εικόνων υψηλής χωρικής διακριτικής ικανότητας*. Πανεπιστήμιο Αιγαίου, Τμήμα Γεωγραφίας.

8. ΠΑΡΑΡΤΗΜΑ

Σύνδεσμος ερωτηματολογίου:

https://forms.office.com/Pages/DesignPageV2.aspx?subpage=design&FormId=7kOJDHDDs0u6UTIfQG8y7P3c-cp-GENNjdfft_NDvo9UMkYwMEI1MTc3SFBHQVFCSjlLN09ONFRNUy4u&Token=ee2774ab31254a8a9c92bf25953a37cc