University of Ioannina

Πληροφορικής και Τηλεπικοινωνιών

Πληροφορική και Δίκτυα

Διπλωματική Εργασία

Μελέτη και υλοποίηση τεχνικών επαύξησης δεδομένων σε προβλήματα ταξινόμησης κειμένου που αφορούν την ασφάλεια υπολογιστών

Στέφανος Μούνγκουλης

Επιβλέπων καθηγητής: Πέτρος Καρβέλης

Αθήνα, Μάιος 2023

Μελέτη και υλοποίηση τεχνικών επαύξησης δεδομένων σε προβλήματα ταξινόμησης κειμένου που αφορούν την ασφάλεια υπολογιστών

Στέφανος Μούνγκουλης

Επιτροπή Επίβλεψης Διπλωματικής Εργασίας

Επιβλέπων Καθηγητής:
Πέτρος Καρβέλης
Επίκουρος καθηγητής & Σχολή Πληροφορικής & Τηλεπικοινωνιών,
Πανεπιστήμιο Ιωαννίνων

Αθήνα, Μάιος 2023

«Θέλω να ευχαριστήσω τους γονείς μου.»

Περίληψη

Κυβερνοαπειλή είναι μία οποιαδήποτε ενέργεια ή συμβάν, που θα μπορούσε να οδηγήσει σε μη επιθυμητό αντίκτυπο στις υποδομές πληροφορικής. Το μέσο κόστος μιας κυβερνοεπίθεσης αποτιμάται περίπου στα 3.8 εκατομμύρια ευρώ. Η αγορά της κυβερνοασφάλειας εκτιμάται ότι θα αγγίξει τα 400 δισεκατομμύρια ευρώ ετησίως, τα αμέσως επόμενα χρόνια. Μόνο το έτος 2021, ανακαλύφθηκαν περισσότερες από 18 χιλιάδες ευπάθειας ασφαλείας. Ορισμένες από τις προσπάθειες οργάνωσης εναντίον του κινδύνου αυτού, αφορούν την δημιουργία των λεγόμενων μοντέλων απειλής που έχουν ως στόχο την προληπτική διαδικασία εντοπισμού των απειλών, προτού αυτές επηρεάσουν κάποιον οργανισμό. Ένα από τα γνωστότερα μοντέλα απειλών είναι το MITRE ATT&CK. Η χειροκίνητη όμως καταγραφή των ευπαθειών από κάποιον αναλυτή ασφαλείας, αποτελεί μία εξαιρετικά χρονοβόρα διαδικασία και έτσι, γίνεται μία προσπάθεια αυτοματοποίησης της διαδικασίας αυτής μέσω υλοποίησης αλγορίθμων επαύξησης δεδομένων κειμένων, προερχόμενων από τον τομέα της μηχανικής μάθησης και επεξεργασίας της φυσικής γλώσσας. Με τα δεδομένα αυτά, να χρησιμοποιούνται στη συνέχεια σε ταξινομητές πολλαπλών ετικετών για σύγκριση και βελτιστοποίηση.

Λέξεις – Κλειδιά

Cyber Security, MITE ATT&CK, Machine Learning, Python, Natural Language Processing

Study and implementation of data augmentation techniques regarding text classification problems that concern computer security

Stefanos Moungkoulis

Abstract

A cyber threat is any action or event that could lead to an unwanted impact on IT infrastructures. The average cost of a cyberattack is estimated around 3.8 million euros. The cyber security market is estimated to reach 400 billion euros per year in the next few years. In the year 2021 alone, more than 18 thousand security vulnerabilities were discovered. Some of the organization's efforts against the risk involves the creation of so-called threat models that aim to proactively identify threats before they affect an organization. One of the best-known threat models is MITRE ATT&CK. However, the manual recording of vulnerabilities by a security analyst is an extremely time-consuming process and thus, an attempt is made to automate this process through the implementation of text data augmentation algorithms, coming from the field of machine learning and natural language processing. With this data later being used in multi-label classifiers for comparison and optimization.

Keywords

Cyber Security, MITRE ATT&CK, Python, Machine Learning, Natural Language Processing

Περιεχόμενα

Κατάλογος Εικόνων / Σχημάτων

Κατάλογος Πινάκων

Συντομογραφίες & Ακρωνύμια

DRM	Digital Rights Management
OWASP	Open Web Application Security
IT	Information Technology
PCI DSS	Payment Card Industry Data Security Standards
GDPR	General Data Protection Regulation
ENISA	European Union Agency for Cybersecurity
NIST	National Institute of Technologies
CVE	Common Vulnerabilities and Exposures
CVSS	Common Vulnerability Scoring System
APT	Advanced Persistent Threats
ICS	Industrial Control Systems
FMX	Fort Meade Experiment
SOC	Security Operations Center
NVD	National Vulnerability Database
NLP	Natural Language Processing
TF-IDF	Term Frequency – Inverse Document Frequency
CAPEC	Common Attack Pattern Enumeration and Classification
BOW	Bag of Words
CBOW	Continuous Bag of Words
SVM	Support Vector Machine

Εισαγωγή

Η χρήση ηλεκτρονικών υπολογιστών ή γενικότερα ηλεκτρονικών συστημάτων αποτελεί πλέον, μέρος της καθημερινότητάς μας. Η συντριπτική πλειοψηφία του πληθυσμού, τουλάχιστον όσον αφορά στις ανεπτυγμένες και στις αναπτυσσόμενες χώρες, έχει τακτική επαφή με την τεχνολογία αυτή, είτε για προσωπική είτε για επαγγελματική χρήση. Τα έσοδα της αγοράς λογισμικού, μόνο στις ΗΠΑ αναμένεται να φθάσουν τα 593 δισεκατομμύρια δολάρια το 2022. Με την κατάσταση αυτή, όπως ήταν αναμενόμενο, προέκυψαν και ζητήματα τα οποία αφορούν την ασφάλεια των συστημάτων αυτών. Η ασφάλεια των υπολογιστών και των δικτύων ή γενικότερα, η ασφάλεια στον κυβερνοχώρο, είναι πλέον, ένα εξαιρετικά κρίσιμο ζήτημα. Η ασφάλεια των συστημάτων, αποτελεί κεντρικό ζήτημα για χώρες, οργανισμούς ακόμα προφανώς και για τους ίδιους τους οικιακούς χρήστες. Οι μηχανισμοί ασφαλείας που επιβάλλονται σε κάθε περίπτωση, έχουν ως στόχο, την διασφάλιση ότι το σύστημα δεν θα βρεθεί ποτέ, σε μία μη επιτρεπόμενη κατάσταση. Μία μη επιτρεπόμενη κατάσταση σε ένα λογισμικό ή λειτουργικό, μπορεί να χαρακτηριστεί και ως ευπάθεια και πιο συγκεκριμένα, μία ευπάθεια, αποτελεί ένα ελάττωμα, ικανό να επιτρέψει σε έναν εισβολέα να αποκτήσει έως και τον πλήρη έλεγχο ενός συστήματος. Τα ελαττώματα αυτά, μπορεί να οφείλονται στον τρόπο με τον οποίο έχει σχεδιαστεί το λογισμικό ή σε κάποιο ελάττωμα στον τρόπο με τον οποίο είναι κωδικοποιημένο. Σύμφωνα με το comparitech (2023) μόνο στο πρώτο τρίμηνο του 2022, δημοσιεύτηκαν περισσότερες από 8 χιλιάδες ευπάθειες λογισμικού. Πρόκειται για αύξηση περίπου 25 τις εκατό, σε σχέση με την ίδια περίοδο του προηγούμενου έτους.

Ο αριθμός αυτός είναι αρκετά μεγάλος και μία προσπάθεια οργάνωσής τους, προκύπτει από την δημιουργία των λεγόμενων μοντέλων απειλών. Τα μοντέλα αυτά, αποτελούν μια προληπτική διαδικασία εντοπισμού υφιστάμενων ή και νέων απειλών. Ένα από τα μοντέλα αυτά, είναι και το MITRE ATT&CK, το οποίο αποτελεί ένα εξαιρετικά λεπτομερές και διασταυρωμένο αποθετήριο πληροφοριών, σχετικά με κακόβουλες ομάδες, τις τεχνικές τους καθώς και τη γνωστή συμπεριφορά τους στις διαδικασίες που χρησιμοποιούν. Αφορά επίσης το λογισμικό καθώς και τα νόμιμα ή και κακόβουλα εργαλεία που αυτές οι ομάδες χρησιμοποιούν. Όμως, η χειροκίνητη καταγραφή και χαρτογράφηση των συνεχόμενα αυξημένων απειλών, αποτελεί μία χρονοβόρα διαδικασία για τους αναλυτές ασφαλείας και έτσι, για την πιο εύκολη λειτουργία των μοντέλων, θα μελετηθούν και θα υλοποιηθούν αλγόριθμοι επαύξησης δεδομένων κειμένων με επεξεργασία φυσικής γλώσσας. Στη συνέχεια, τα δεδομένα αυτά θα χρησιμοποιούνται σε ταξινομητές πολλαπλών ετικετών για

σύγκριση και βελτιστοποίηση ούτως ώστε η κατηγοριοποίηση να γίνεται με όσο το δυνατόν πιο αυτοματοποιημένο τρόπο.

1. Ασφάλεια υπολογιστών

Η χρήση των υπολογιστών είναι πλέον απαραίτητη για μία σύγχρονη κοινωνία και η εξάρτηση αυτή, συνεχώς αυξάνεται. Οργανισμοί, τόσο διαφορετικοί όπως εκκλησίες, επιχειρήσεις, εταιρείες, στρατός και κράτη, επωφελούνται από την δύναμη και την ευκολία χρήσης των συστημάτων υπολογιστών. Η παγκόσμια και ευρεία χρήση δικτύων, η σύνδεση των συστημάτων που κυμαίνονται από προσωπικούς υπολογιστές έως και μεγάλα mainframe συστήματα, διευκολύνει την ανταλλαγή, την μεταφορά και την επικοινωνία δεδομένων και έχει φέρει τον κόσμο κοντά. Αυτή η παγκόσμια εμβέλεια ορισμένων δικτύων, καθιστά την μεταφορά δεδομένων σε χώρες μακρινές, τόσο εύκολη όσο μία πρόσβαση σε απλή τηλεφωνική γραμμή και ένα τερματικό.

Τα δεδομένα σε ορισμένα υπολογιστικά συστήματα είναι τόσο σημαντικά, που μη εξουσιοδοτημένη πρόσβαση σε αυτά ή και καταστροφή τους θα μπορούσε να έχει καταστροφικές οικονομικές και όχι μόνο, συνέπειες. Η ανάγκη προστασίας των δεδομένων υπολογιστή από μη εξουσιοδοτημένη πρόσβαση, δεν είναι κάτι το καινούργιο. Ωστόσο, τα τελευταία χρόνια, έχει γίνει πιο σημαντικό λόγω της σχεδόν απόλυτης εξάρτησης των οργανισμών, από τα δεδομένα τους.

Η ασφάλεια υπολογιστών, περιλαμβάνει έννοιες και μεθόδους που αφορούν την προστασία ευαίσθητων πόρων σε συστήματα υπολογιστών. Η αρχή, γίνεται με την εφαρμογή πολιτικών που ρυθμίζουν την πρόσβαση στους λεγόμενους προστατευόμενους πόρους. Στην τεχνολογία, η έμφαση δίνεται στους εκάστοτε μηχανισμούς επιβολής αυτών των πολιτικών. Η ασφάλεια των υπολογιστών αποκτά σημασία, όταν τα συστήματα υπολογιστών, επεξεργάζονται ευαίσθητα δεδομένα ή εκτελούν ευαίσθητες υπηρεσίες και μπορεί να χωριστεί σε τέσσερα μέρη.

- Έλεγχο της πρόσβασης στο ίδιο το σύστημα υπολογιστών.
- Έλεγχο της πρόσβασης στους πόρους που διαχειρίζεται το σύστημα.
- Προστασία των δεδομένων κατά την μεταφορά τους μεταξύ των συστημάτων.
- Διασφάλιση των εφαρμογών έναντι κακόβουλων εισροών.

Στην ασφάλεια των υπολογιστών, σύμφωνα με τον Gollmann (2011, pp. 544-545) μπορεί κανείς να παρακολουθήσει την εξέλιξη από τα αυτόνομα συστήματα που επιβάλλουν πολιτικές επικεντρωμένες στον χρήστη, σε επίπεδο λειτουργικού συστήματος έως τα κατανεμημένα συστήματα, που επιβάλλουν πολιτικές επικεντρωμένες στον κώδικα σε

πρόγραμμα περιήγησης στο διαδίκτυο ή ακόμα και σε ιστοσελίδες. Οι τάσεις αυτές, καθοδηγούνται από τις αλλαγές στις αρχιτεκτονικές του IT και συμβαδίζουν με τις αλλαγές της. Συναφή ζητήματα, αποτελούν η διαχείριση των ψηφιακών δικαιωμάτων (DRM), η αξιόπιστη πληροφορική, η μετάβαση από τις πολιτικές για έναν μόνο οργανισμό, στις πολιτικές σε ομοσπονδιακά συστήματα και από την αξιολόγηση της ασφάλειας των αυτόνομων συστημάτων, στις προκλήσεις που θέτει η αξιολόγηση της ασφάλειας των σημερινών επεκτάσιμων συστημάτων λογισμικού.

Εικόνα 1: Ασφάλεια Υπολογιστών

Η ασφάλεια των υπολογιστών έχει τις ρίζες της στη δεκαετία του 60. Για παράδειγμα, οι δακτύλιοι προστασίας εισήχθησαν για να επιλύσουν το πρόβλημα προστασίας των δεδομένων του χρήστη και του συστήματος κατά την διάρκεια εκτέλεσης διεργασιών στο λειτουργικό σύστημα Multics, ένα ζήτημα που έπρεπε να αντιμετωπιστεί στα αναδυόμενα συστήματα πολλαπλών χρηστών.

Στην περεταίρω εξέλιξη της ασφάλειας υπολογιστών, δύο σκέλη της ανάπτυξης είναι αλληλένδετα. Οι πολιτικές ασφάλειας που καλείται να επιβάλει ένα σύστημα υπολογιστών και το αρχιτεκτονικό επίπεδο του συστήματος υπολογιστών όπου επιβάλλεται η κάθε πολιτική. Στην ασφάλεια υπολογιστών, ο έλεγχος πρόσβασης που εμποδίζει τη μη εξουσιοδοτημένη πρόσβαση σε συστήματα και πόρους είναι η κύρια υπηρεσίας ασφαλείας για την επίτευξη της εμπιστευτικότητας, δηλαδή την αποτροπή της μη εξουσιοδοτημένης αποκάλυψης πληροφοριών και της ακεραιότητας, δηλαδή την αποτροπή της μη εξουσιοδοτημένης τροποποίησης πληροφοριών. Οι μηχανισμοί ελέγχουν πρόσβασης,

καθορίζουν τον τρόπο με τον οποίο ένα υποκείμενο (διεργασία) μπορεί να έχει πρόσβαση σε έναν πόρο (αντικείμενο). Ο έλεγχος πρόσβασης, μπορεί να χωριστεί σε τρεις κύριες εργασίες.

- Στη ρύθμιση της πολιτικής ασφαλείας, στην οποία οι εντολείς έχουν εξουσιοδότηση πρόσβασης σε ορισμένους μόνο πόρους.
- Στη πιστοποίηση της γνησιότητας των αποδεικτικών στοιχείων που παρέχονται, με αίτηση πρόσβασης.
- Στην αξιολόγηση της αίτησης, σε σχέση με τη δεδομένη πολιτική.

Οι δύο πολιτικές που σχετίζονται με τα πρώτα στάδια της ασφάλειας των υπολογιστών είναι γνωστές ως διακριτικός και υποχρεωτικός έλεγχος πρόσβασης. Σε ένα λειτουργικό σύστημα, ο έλεγχος πρόσβασης, αποφασίζει εάν μία διεργασία που εκτελείται υπό μία δεδομένη ταυτότητα διακομιστή, επιτρέπεται να έχει πρόσβαση σε έναν δεδομένο πόρο. Γενικότερα, μία πολιτική, καθορίζει το τι επιτρέπεται να κάνουν οι εντολείς. Ο έλεγχος ταυτότητας δημιουργεί μία διαδικασία η οποία εκτελείται κάτω από την ταυτότητα του χρήστη που επαληθεύτηκε κατά τη σύνδεση με συνηθισμένο πρότυπο, τον έλεγχο, βάσει κωδικούς πρόσβασης. Τέλος, υπάρχει και η απόφαση πολιτικής με την οποία, γίνεται ο έλεγχος για την συμμόρφωση του ελέγχου πρόσβασης.

1.1 Προγράμματα λογισμικού

Λογισμικό, αποτελεί ένα σύνολο οδηγιών, δεδομένων και προγραμμάτων που χρησιμοποιούνται για την λειτουργία των υπολογιστών και την εκτέλεση ορισμένων εργασιών. Είναι το αντίθετο του υλικού, το οποίο περιγράφει τις φυσικές πτυχές ενός υπολογιστή. Το λογισμικό είναι ένας γενικός όρος που χρησιμοποιείται για να αναφερθεί στις εφαρμογές, στα σενάρια και τα προγράμματα που εκτελούνται σε μία οποιαδήποτε συσκευή. Μπορεί επίσης να θεωρηθεί, ως το μεταβλητό μέρος ενός υπολογιστή, ενώ το υλικό, θεωρείται ως το αμετάβλητο μέρος.

Οι δύο κύριες κατηγορίες λογισμικού είναι το λογισμικό εφαρμογών και το λογισμικό συστήματος. Κάθε εφαρμογή, αποτελεί ένα λογισμικό που ικανοποιεί μία συγκεκριμένη ανάγκη ή την εκτέλεση μίας εργασίας. Οι εφαρμογές, είναι ο πιο συνηθισμένος τύπος λογισμικού και αποτελεί ένα πακέτο λογισμικού που εκτελεί μία συγκεκριμένη λειτουργία

για έναν χρήστη ή, σε ορισμένες περιπτώσεις, για μία άλλη εφαρμογή. Μία εφαρμογή, μπορεί να είναι αυτοτελής ή μπορεί να είναι μία ομάδα προγραμμάτων που εκτελούν την εφαρμογή για τον χρήστη. Κάποια παραδείγματα σύγχρονων εφαρμογών περιλαμβάνουν

- Σουίτες γραφείου (Microsoft Office, Open Office)
- Λογισμικό γραφικών (Adobe Photoshop, Gimp)
- Βάσεις δεδομένων
- Προγράμματα περιήγησης στο διαδίκτυο
- Επεξεργαστές κειμένου
- Προγράμματα επεξεργασίας εικόνας
- Πλατφόρμες επικοινωνίας

Το λογισμικό συστήματος από την άλλη, έχει σχεδιαστεί κατ' αυτόν τον τρόπο ώστε να λειτουργεί το υλικό ενός υπολογιστή και παρέχει μία πλατφόρμα, για την εκτέλεση άλλων εφαρμογών πάνω σε αυτή. Το είδος του λογισμικού αυτού, συντονίζει τις δραστηριότητες και τις λειτουργίες του υλικού και του λογισμικού. Επιπλέον, ελέγχει τις λειτουργίες του υλικού του υπολογιστή και παρέχει ένα περιβάλλον ή μία πλατφόρμα για την εργασία όλων των άλλων τύπων λογισμικού. Το λειτουργικό σύστημα (Microsoft Windows, Linux) είναι το καλύτερο παράδειγμα λογισμικού συστήματος. Άλλα παραδείγματα, αποτελούν το υλικολογισμικό (firmware), οι μεταφραστές γλωσσών υπολογιστών και τα βοηθητικά προγράμματα συστήματος.

Άλλοι τύποι λογισμικού, περιλαμβάνουν το λογισμικό προγραμματισμού που παρέχει τα εργαλεία προγραμματισμού τα οποία απαιτούνται από τους προγραμματιστές λογισμικού, ως ενδιάμεσο λογισμικό, το οποίο βρίσκεται μεταξύ του λογισμικού συστήματος και των εφαρμογών και επίσης, υπάρχει ο τύπος του λογισμικού οδηγών το οποίο χρησιμοποιείται από τις συσκευές και τα περιφερειακά του υπολογιστή.

Το πρώιμο λογισμικό, είχε γραφτεί για συγκεκριμένους υπολογιστές και πωλούταν μαζί με το υλικό στο οποίο προβλεπόταν να τρέχει. Στη δεκαετία του 1980, το λογισμικό ξεκίνησε να πωλείται σε δισκέτες και αργότερα σε CD Και σε DVD. Σήμερα, το μεγαλύτερο μέρος του διαθέσιμου λογισμικού αγοράζεται και αποκτάται με την χρήση του διαδικτύου. Το λογισμικό, μπορεί να βρεθεί σε δικτυακούς τόπους πωλητών ή και σε δικτυακούς τόπους παρόχων υπηρεσιών εφαρμογών.

1.2 Ευπάθεια ασφαλείας

Ευπάθεια ασφαλείας, σύμφωνα με το Rapid7 (2023) ορίζεται ως μία αδυναμία στην υπολογιστική λογική, για παράδειγμα κώδικας, που βρίσκεται σε στοιχεία λογισμικού ή και υλικού που, όταν γίνεται εκμετάλλευση, έχει ως αποτέλεσμα ένα αρνητικό αντίκτυπο στην εμπιστευτικότητα, την ακεραιότητα ή και τη διαθεσιμότητα. Ο μετριασμός των τρωτών σημείων σε αυτό το πλαίσιο, συνήθως περιλαμβάνει αλλαγές κωδικοποίησης, αλλά θα μπορούσε επίσης να περιλαμβάνει αλλαγές ή ακόμα και καταργήσεις προδιαγραφών. Γενικότερα μιλώντας, μία ευπάθεια μπορεί να είναι οτιδήποτε μπορεί να απειλήσει την ασφάλεια ενός υπολογιστικού συστήματος.

Οι ευπάθειες, μπορούν να οδηγήσουν σε διαρροές δεδομένων και τελικά, σε παραβιάσεις αυτών. Μία διαρροή δεδομένων, σημαίνει ότι τα δεδομένα διαρρέουν καταλάθος μέσα από έναν οργανισμό, σε αντίθεση με την παραβίαση δεδομένων, η οποία είναι αποτέλεσμα κλοπής δεδομένων. Η διαρροή δεδομένων, είναι συνήθως αποτέλεσμα ενός λάθους, πολύ συχνά ανθρώπινου όπως για παράδειγμα η αποστολή ενός εγγράφου με ευαίσθητες ή εμπιστευτικές πληροφορίες σε λάθος παραλήπτη ηλεκτρονικού ταχυδρομείου, αποθήκευση δεδομένων σε δημόσια κοινή χρήση ή ύπαρξη δεδομένων σε μία ξεκλείδωτη συσκευή σε δημόσιο χώρο, προσβάσιμη σε τρίτους. Μερικά συχνά παραδείγματα ευπαθειών αφορούν τα εξής.

- Μη επιδιορθωμένο λογισμικό
- Λανθασμένη διαμόρφωση συστήματος
- Αδύναμα διαπιστευτήρια
- Δραστηριότητες phishing
- Διαμορφώσεις αξιοπιστίας
- Παραβιασμένα διαπιστευτήρια
- Κακόβουλος insider
- Ανυπαρξία ή αδυναμίες κρυπτογράφησης
- Άγνωστες και zero day μέθοδοι

Το μη επιδιορθωμένο λογισμικό, είναι ένας κώδικας υπολογιστή που περιέχει γνωστές αδυναμίες ασφαλείας. Τα μη επιδιορθωμένα τρωτά σημεία, αναφέρονται σε εκείνες τις αδυναμίες που επιτρέπουν στους εισβολείς να αξιοποιήσουν το σφάλμα ασφαλείας εκτελώντας κακόβουλο κώδικα.

Οι εσφαλμένες διαμορφώσεις ασφαλείας είναι στοιχεία ελέγχου ασφαλείας τα οποία δεν έχουν ρυθμιστεί σωστά ή αφήνονται ανασφαλή, θέτοντας τα συστήματα και τα δεδομένα σε κίνδυνο. Οι εσφαλμένες διαμορφώσεις αυτές είναι δυνατό να προκύψουν και ύστερα από κάποιο τεχνικό πρόβλημα.

Ένας επιτιθέμενος, μπορεί να χρησιμοποιήσει διαφορετικού είδους επιθέσεις έτσι ώστε να "μαντέψει" τυχόν αδύναμους κωδικούς, με την διαδικασία αυτή να αποφέρει πλήρη πρόσβαση στα συστήματα ενός δικτύου. Οι επιθέσεις phishing, προέρχονται από απατεώνες μεταμφιεσμένους ως αξιόπιστες πηγές και μπορούν να διευκολύνουν σε κακόβουλους χρήστες την πρόσβαση, σε όλους τους τύπους ευαίσθητων δεδομένων με τον πιο συνηθισμένο τρόπο, να αποτελεί το email phishing.

Επίσης, οι εισβολείς μπορούν να εκμεταλλευτούν διαμορφώσεις αξιοπιστίας που έχουν ρυθμιστεί να επιτρέπουν ή να απλοποιούν την πρόσβαση μεταξύ συστημάτων για παράδειγμα τοποθετημένες μονάδες δίσκου ή απομακρυσμένες υπηρεσίες. Με την απόκτηση πρόσβασης σε ένα σύστημα, ο επιτιθέμενος μπορεί να προχωρήσει στην παραβίαση και των υπολοίπων που "εμπιστεύονται" το αρχικά παραβιασμένο σύστημα. Επιπλέον, ένας εισβολέας μπορεί να εκμεταλλευτεί τυχόν παραβιασμένα διαπιστευτήρια για να αποκτήσει μη εξουσιοδοτημένη πρόσβαση σε ένα σύστημα του δικτύου. Τα παραβιασμένα διαπιστευτήρια είναι πιθανό να εκμεταλλευτούν σε τυχόν διαφορετικά συστήματα, σε περίπτωση που έχει υπάρξει επαναχρησιμοποίησή τους.

Σημαντικό ζήτημα ασφαλείας, αποτελεί κάποιος υπάλληλος ή ένας πωλητής, που μπορεί να έχει πρόσβαση σε κρίσιμα συστήματα εάν αποφασίσει να εκμεταλλευτεί την πρόσβαση αυτή για να υποκλέψει ή να καταστρέψει κάποιες πληροφορίες. Αυτό, αποκτά ιδιαίτερη σοβαρότητα, ιδίως όταν μιλάμε για κάποιους χρήστες που έχουν πρόσβαση σε πολλαπλά κρίσιμα συστήματα, όπως για παράδειγμα κάποιοι administrators.

Με την εκμετάλλευση ζητημάτων κρυπτογράφησης, ένας επιτιθέμενος μπορεί να υποκλέψει την επικοινωνία μεταξύ συστημάτων στο δίκτυο και να κλέψει ευαίσθητες πληροφορίες. Οι zero day ευπάθειες είναι κάποιες συγκεκριμένες ευπάθειες ασφαλείας λογισμικού που είναι γνωστές στους επιτιθέμενους, για τις οποίες δεν έχει υπάρξει ακόμα κάποια επιδιόρθωση και ίσως κάποιες φορές ο προμηθευτής, να αγνοεί την ύπαρξη της συγκεκριμένης ευπάθειας.

1.3 Cybersecurity Frameworks

Ένα πλαίσιο κυβερνοασφάλειας είναι ένα σύνολο κατευθυντήριων γραμμών, βέλτιστων πρακτικών και προτύπων που έχουν σχεδιαστεί για να βοηθούν τους οργανισμούς να διαχειρίζονται και να μειώνουν τους κινδύνους κυβερνοασφάλειας. Σκοπός ενός πλαισίου κυβερνοασφάλειας, είναι να παρέχει μια δομημένη προσέγγιση για την κυβερνοασφάλεια που μπορεί να προσαρμοστεί ώστε να ανταποκρίνεται στις συγκεκριμένες ανάγκες και απαιτήσεις ενός οργανισμού. Ένα τυπικό πλαίσιο κυβερνοασφάλειας, περιλαμβάνει ένα σύνολο ελέγχων ασφαλείας, πολιτικών και διαδικασιών που αφορούν διάφορες πτυχές της κυβερνοασφάλειας, όπως η διαχείριση κινδύνων, ο έλεγχος πρόσβασης, η αντιμετώπιση περιστατικών και η προστασία δεδομένων. Το πλαίσιο, μπορεί επίσης να περιλαμβάνει εργαλεία και τεχνολογίες για την εφαρμογή και την παρακολούθηση των ελέγχων ασφαλείας, καθώς και καθοδήγηση σχετικά με τον τρόπο αξιολόγησης και βελτίωσης της κατάστασης κυβερνοασφάλειας κάποιου οργανισμού.

Πολλά πλαίσια κυβερνοασφάλειας, αναπτύσσονται από βιομηχανικές ομάδες, κυβερνητικούς οργανισμούς ή διεθνείς οργανισμούς και είναι σχεδιασμένα ώστε να μπορούν να προσαρμοστούν σε ένα ευρύ φάσμα οργανισμών και κλάδων. Με την εφαρμογή ενός πλαισίου κυβερνοασφάλειας, οι οργανισμοί μπορούν να βελτιώσουν την συνολική τους θέση στον κυβερνοχώρο, να μειώσουν τον κίνδυνο επιθέσεων καθώς επίσης και να προστατεύσουν καλύτερα τις ευαίσθητες πληροφορίες και τα περιουσιακά τους στοιχεία. Τα πλαίσια κυβερνοασφάλειας είναι συχνά υποχρεωτικά ή τουλάχιστον ενθαρρύνονται έντονα, προς τις εταιρείες εκείνες οι οποίες θέλουν να συμμορφωθούν με κρατικούς, βιομηχανικούς και διεθνείς κανονισμούς κυβερνοασφάλειας. Για παράδειγμα, προκειμένου μία εταιρεία να μπορεί να διαχειρίζεται συναλλαγές με πιστωτικές κάρτες, θα πρέπει πρώτα να περάσει από έλεγχο που πιστοποιεί τη συμμόρφωσή της με το πλαίσιο Payment Card Industry Data Security Standards (PCI DSS).

Πιο συγκεκριμένα, πολλά πλαίσια κυβερνοασφάλειας δίνουν μεγάλη έμφαση στη διαχείριση κινδύνων, η οποία περιλαμβάνει τον εντοπισμό και την αξιολόγηση των κινδύνων κυβερνοασφάλειας και την εφαρμογή ελέγχων για τον μετριασμό αυτών των κινδύνων. Αυτό μπορεί να περιλαμβάνει τη διενέργεια τακτικών αξιολογήσεων κινδύνου, την ανάπτυξη σχεδίων διαχείρισης κινδύνου και την εφαρμογή ελέγχων ασφαλείας με βάση το επίπεδο κινδύνου. Ορισμένα άλλα πλαίσια κυβερνοασφάλειας, έχουν σχεδιαστεί για να βοηθήσουν τους οργανισμούς να συμμορφωθούν με τις νομικές και κανονιστικές απαιτήσεις που σχετίζονται με την κυβερνοασφάλεια, όπως ο Γενικός Κανονισμός της ΕΕ

για την Προστασία Δεδομένων (GDPR) ή ο νόμος των ΗΠΑ για τον εκσυγχρονισμό της ασφάλειας των πληροφοριών (FISMA). Η συμμόρφωση με αυτά τα πλαίσια μπορεί να βοηθήσει τους οργανισμούς να αποφύγουν κυρώσεις και άλλες νομικές συνέπειες που σχετίζονται με παραβιάσεις της ασφάλειας στον κυβερνοχώρο.

Ένα από τα βασικά πλεονεκτήματα των πλαισίων κυβερνοασφάλειας είναι ότι είναι γενικά σχεδιασμένα ώστε να είναι ευέλικτα και προσαρμόσιμα σε διαφορετικούς οργανισμούς και κλάδους. Αυτό σημαίνει ότι οι οργανισμοί μπορούν να προσαρμόσουν το πλαίσιο ώστε να ανταποκρίνεται στις συγκεκριμένες ανάγκες και απαιτήσεις τους και μπορούν να επιλέξουν να εφαρμόσουν μόνο τους ελέγχους και τις πολιτικές που είναι πιο σχετικές με τις δραστηριότητές τους. Επίσης, πολλά πλαίσια κυβερνοασφάλειας έχουν σχεδιαστεί για να βελτιώνονται και να ενημερώνονται συνεχώς με βάση τις νέες απειλές και τους κινδύνους. Αυτό μπορεί να περιλαμβάνει τη διενέργεια τακτικών αξιολογήσεων ασφαλείας, την αναθεώρηση και επικαιροποίηση πολιτικών και διαδικασιών και την εφαρμογή νέων τεχνολογιών και εργαλείων ασφαλείας. Τέλος, όσον αφορά την ενσωμάτωση τα πλαίσια κυβερνοασφάλειας μπορούν να βοηθήσουν τους οργανισμούς να ενσωματώσουν διάφορες πτυχές του προγράμματος κυβερνοασφάλειας, όπως η διαχείριση κινδύνων, η αντιμετώπιση περιστατικών και η εκπαίδευση ευαισθητοποίησης σε θέματα ασφάλειας. Παρέχοντας ένα ενιαίο πλαίσιο για τη διαχείριση της ασφάλειας στον κυβερνοχώρο, οι οργανισμοί μπορούν να διασφαλίσουν ότι όλες οι πτυχές του προγράμματός τους είναι ευθυγραμμισμένες και συνεργάζονται αποτελεσματικά.

Η λίστα με τα πιο γνωστά πλαίσια κυβερνοασφάλειας περιλαμβάνει

- Cyber Essentials
- Center for Internet Security (CIS) Framework
- International Office of Standardization (ISO) 27001
- American Institute of CPAs' Service Organization Control (SOC 2)
- Control Objectives for Information Technology (COBIT)
- Payment Card Industry Data Security Standard (PCI DSS)
- Cloud Security Alliance (CSA)
- Cloud Controls Matrix (CCM)
- European Union Agency for Cybersecurity (ENISA) National Capabilities Assessment Framework
- National Institute of Technologies (NIST) Cybersecurity Framework (CSF)
- ISO/IEC 27001

- SANS Top 20
- Factor Analysis of Information Risk (FAIR)
- Cybersecurity Information Sharing Act (CISA)
- Health Insurance Portability and Accountability Act (HIPAA)
- IEC 62443

1.3.1 NIST

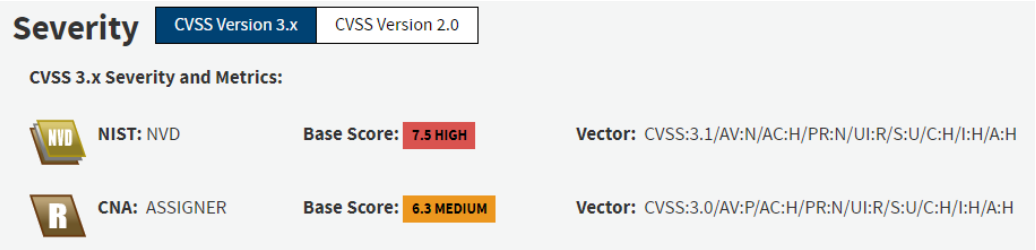
Το Εθνικό Ινστιτούτο Προτύπων και Τεχνολογίας (NIST) ιδρύθηκε το 1901 και αποτελεί σήμερα, μέρος του Υπουργείου Εμπορίου των ΗΠΑ. Το NIST, είναι ένα από τα παλαιότερα εργαστήρια φυσικών επιστημών των ΗΠΑ. Όσον αφορά τον τομέα της κυβερνοασφάλειας, το NIST αναπτύσσει πρότυπα κυβερνοασφάλειας, δίνει κατευθυντήριες γραμμές, βέλτιστες πρακτικές και άλλους πόρους, για να καλύψει τις ανάγκες της αμερικανικής βιομηχανίας, των ομοσπονδιακών υπηρεσιών και του ευρύτερου κοινού. Οι δραστηριότητες, κυμαίνονται από την παραγωγή συγκεκριμένων πληροφοριών με τις οποίες οι οργανισμοί, μπορούν να εφαρμόσουν άμεσα στην πράξη έως την πιο μακροπρόθεσμη έρευνα, που προβλέπει την πρόοδο των τεχνολογιών και τις μελλοντικές προκλήσεις. Το NIST, προάγει την κατανόηση και βελτιώνει την διαχείριση των κινδύνων για την προστασία της ιδιωτικής ζωής, ορισμένοι από τους οποίους σχετίζονται άμεσα με την ασφάλεια στον κυβερνοχώρο. Οι τομείς προτεραιότητας, περιλαμβάνουν την κρυπτογραφία, την εκπαίδευση και το εργατικό δυναμικό, τις αναδυόμενες τεχνολογίες, τη διαχείριση κινδύνων, τη διαχείριση ταυτότητας και πρόσβασης, τις μετρήσεις, την προστασία της ιδιωτικής ζωής και τέλος, τα αξιόπιστα δίκτυα και τις αξιόπιστες πλατφόρμες.

Σχήμα 1: NIST logo

Στο NIST, περιλαμβάνεται λίστα που ανανεώνεται συνεχώς περιλαμβάνοντας νέες ευπάθειες που ανιχνεύονται καθημερινά. Στις ευπάθειες αυτές, αποδίδεται ένα αναγνωριστικό CVE. Ο τρόπος που ορίζεται μία ευπάθεια από το CVE έχει ως εξής.

“Μία αδυναμία στην υπολογιστική λογική, για παράδειγμα κώδικας, που εντοπίζεται σε στοιχεία λογισμικού ή και υλικού, η οποία, αν αξιοποιηθεί, έχει ένα αρνητικό αποτέλεσμα στην εμπιστευτικότητα, την ακεραιότητα ή και τη διαθεσιμότητα ενός προγράμματος ή συστήματος. Η επίλυση της ευπάθειας αυτής, περιλαμβάνει συνήθως αλλαγές στον κώδικά της αλλά επίσης και αλλαγές στις προδιαγραφές ή ακόμη και απαλοιφές προδιαγραφών, για παράδειγμα, κατάργηση των επηρεαζόμενων πρωτοκόλλων ή λειτουργιών στο σύνολό τους”.

Για την παροχή ενός ποιοτικού μέτρου σοβαρότητας σύμφωνα με το website του NIST, χρησιμοποιείται το σύστημα Common Vulnerability Scoring System (CVSS). Η βαθμολογία αυτή, προκύπτει με βάση τις μετρήσεις εκμεταλλευσιμότητας της ευπάθειας καθώς και των πιθανών επιπτώσεών της. Το εύρος της βαθμολογίας κυμαίνεται από το 1 έως το 10 και, όσο υψηλότερος είναι αυτός ο αριθμός, τόσο υψηλότερη είναι και η σοβαρότητα της ευπάθειας. Ένα παράδειγμα, φαίνετραι στην Εικόνα 2.



Εικόνα 2: Σύστημα CVSS

1.3.2 ENISA

Ο οργανισμός της Ευρωπαϊκής Ένωσης που αφορά την κυβερνοασφάλεια ονομάζεται ENISA. Σύμφωνα με το website της ENISA (2023), αποτελεί τον οργανισμό αυτόν ο οποίος είναι αφιερωμένος στην επίτευξη ενός υψηλού και κοινού επιπέδου κυβερνοασφάλειας σε ολόκληρη την ευρωπαϊκή ήπειρο. Ιδρυμένος το 2004 και ενισχυμένος από την πρωτοβουλία EU Cybersecurity Act ο οργανισμός της Ευρωπαϊκής Ένωσης για την ασφάλεια στον κυβερνοχώρο, συμβάλλει στην πολιτική της ΕΕ στον κυβερνοχώρο, ενισχύει την αξιοπιστία των προϊόντων, υπηρεσιών και διαδικασιών, συνεργάζεται με τα κράτη μέλη και φορείς της ΕΕ και βοηθά την Ευρώπη να προετοιμαστεί για τις προκλήσεις στον κυβερνοχώρο. Μέσω της ανταλλαγής γνώσεων, της δημιουργίας ικανοτήτων και της ευαισθητοποίησης, ο Οργανισμός συνεργάζεται με τους βασικούς ενδιαφερόμενους φορείς για να ενισχύσει την

εμπιστοσύνη στη συνδεδεμένη οικονομία, να ενισχύσει την ανθεκτικότητα των υποδομών της Ένωσης και, τελικά, να διατηρήσει την ευρωπαϊκή κοινωνία και τους πολίτες, ψηφιακά ασφαλείς.

Παρέχει ένα πολύ ευρύ φάσμα υπηρεσιών και πόρων όπως:

- Ανάπτυξη ικανοτήτων στον τομέα της κυβερνοασφάλειας, παρέχοντας κατάρτιση και καθοδήγηση στα κράτη μέλη για να βοηθήσει να αναπτύξουν τις ικανότητές τους στον τομέας της κυβερνοασφάλειας.
- Πληροφορίες και ανάλυση απειλών, παρέχοντας εμπειρογνωμοσύνη και πόρους για να βοηθήσει τα κράτη μέλη, να παρακολουθούν και να ανταποκρίνονται σε απειλές και περιστατικά κυβερνοασφάλειας.
- Πρότυπα και βέλτιστες πρακτικές, εργαζόμενοι στην ανάπτυξη και την προώθηση προτύπων και βέλτιστων πρακτικών για την ασφάλεια στον κυβερνοχώρο.
- Έρευνα για την ασφάλεια, διεξάγοντας ανάλυση και υποστήριξη στην ανάπτυξη αποτελεσματικών πολιτικών και στρατηγικών που αφορούν την ασφάλεια στον κυβερνοχώρο.
- Πρωτοβουλίες ευαισθητοποίησης, διεξάγοντας εκστρατείες για την ευαισθητοποίηση του κοινού σε θέματα κυβερνοασφάλειας και την προώθηση βέλτιστων πρακτικών για την κυβερνοασφάλεια.

Σχήμα 2: European Union Agency for Cybersecurity

Η πανδημία COVID-19 έχει τονίσει την ανάγκη για περισσότερη ασφάλεια στον ψηφιακό κόσμο. Οι πολίτες της Ευρωπαϊκής Ένωσης, έχουν αυξήσει την παρουσία τους στο διαδίκτυο τόσο για τις προσωπικές τους ανάγκες όσο και για τις επαγγελματικές. Οι εγκληματίες του κυβερνοχώρου, εκμεταλλεύτηκαν αυτή την κατάσταση, στοχεύοντας κυρίως τις επιχειρήσεις ηλεκτρονικού εμπορίου και ηλεκτρονικών πληρωμών, καθώς και διάφορα συστήματα υγειονομικής περίθαλψης. Το όραμα της ENISA, αφορά μία αξιόπιστη και ασφαλή στον κυβερνοχώρο Ευρώπη σε συνεργασία με την ευρύτερη κοινότητα.

1.3.3 **PCI DSS**

Πρόκειται για ένα σύνολο προτύπων ασφαλείας, που αναπτύχθηκαν από το Συμβούλιο Προτύπων Ασφαλείας της Βιομηχανίας Καρτών Πληρωμών για να βοηθήσουν τους οργανισμούς που δέχονται πληρωμές με πιστωτικές και χρεωστικές κάρτες να προστατεύσουν το απόρρητο και την ασφάλεια των δεδομένων των κατόχων καρτών. Τα πρότυπα, ισχύουν για όλους εκείνους τους οργανισμούς που αποθηκεύουν, επεξεργάζονται ή διαβιβάζουν δεδομένα κατόχων καρτών, ανεξάρτητα από το μέγεθος ή τον όγκο των συναλλαγών τους. Αποτελείται από ένα σύνολο 12 απαιτήσεων, οι οποίες οργανώνονται σε έξι κατηγορίες. Οι απαιτήσεις αυτές, καλύπτουν ένα ευρύ φάσμα μέτρων ασφαλείας, συμπεριλαμβανομένης της ασφάλειας δικτύου, του ελέγχου πρόσβασης, της κρυπτογράφησης και της διαχείρισης ευπαθειών. Συγκεκριμένα, οι απαιτήσεις αυτές είναι.

- Κατασκευή αλλά και διατήρηση ασφαλούς δικτύου.
- Προστασία των δεδομένων των κατόχων καρτών.
- Συντήρηση ενός προγράμματος διαχείρισης ευπαθειών.
- Εφαρμογή ισχυρών μέτρων ελέγχου πρόσβασης.
- Τακτική παρακολούθηση και δοκιμή των δικτύων.
- Διατήρηση πολιτικής ασφάλειας πληροφοριών.

Η συμμόρφωση με το PCI DSS είναι υποχρεωτική για όλους τους οργανισμούς που διαχειρίζονται πληρωμές με χρεωστικές ή πιστωτικές κάρτες και η μη συμμόρφωση με το πρότυπο αυτό, μπορεί να οδηγήσει σε σημαντικές οικονομικές και νομικές κυρώσεις έχοντας ως συνέπεια την ζημιά στην φήμη του οργανισμού.

1.3.4 OWASP

Το Open Web Application Security Project είναι ένα μη κερδοσκοπικό ίδρυμα που εργάζεται για την βελτίωση της ασφάλειας του λογισμικού. Μέσω έργων λογισμικού ανοικτού κώδικα, υπό την ηγεσία μελών της κοινότητας, των εκατοντάδων τοπικών τμημάτων παγκοσμίως και των δεκάδων χιλιάδων μελών και κορυφαίων εκπαιδευτικών και επιμορφωτικών συνεδρίων, το ίδρυμα OWASP είναι η πηγή για προγραμματιστές και τεχνολόγους όσον αφορά την ασφάλεια του διαδικτύου.

Το OWASP Top 10 είναι ένα αναφορικό έγγραφο, που περιγράφει τις 10 πιο κρίσιμες ανησυχίες που αφορούν την ασφάλεια εφαρμογών ιστού. Σύμφωνα με το website του

OWASP η έκθεση, συντάσσεται από ομάδα εμπειρογνωμόνων ασφαλείας από όλο τον κόσμο και τα δεδομένα, προέρχονται από διάφορους οργανισμούς και στη συνέχεια αναλύονται. Το Top 10 χαρακτηρίζεται ως ένα "έγγραφο ευαισθητοποίησης" και συνίσταται σε όλους τους οργανισμούς, να το ενσωματώσουν στις διαδικασίες τους.

Εικόνα 3: Open Web Application Security Project

Η δική τους λίστα που αφορά τις δέκα πιο κρίσιμες ανησυχίες του έτους 2021 έχει ως εξής:

1) Παραβιασμένος έλεγχος πρόσβασης
2) Αστοχίες κρυπτογράφησης
3) Injections
4) Μη ασφαλής σχεδιασμός
5) Λανθασμένες ρυθμίσεις ασφαλείας
6) Ευάλωτα και μη αναβαθμισμένα στοιχεία
7) Αποτυχίες ταυτοποίησης και πιστοποίησης
8) Αποτυχίες λογισμικού και ακεραιότητας δεδομένων
9) Αποτυχίες καταγραφής και παρακολούθησης της ασφάλειας
10) Πλαστογραφήσεις αιτημάτων από την πλευρά του διακομιστή

1.4 Μοντελοποίηση απειλών

Η μοντελοποίηση απειλών, σύμφωνα με το Canadian Center for Cyber Security (2019) είναι μία προληπτική διαδικασία εντοπισμού των κινδύνων και των απειλών, που είναι πιθανό να επηρεάσουν έναν οργανισμό και στη συνέχεια, ο σχεδιασμός και η εφαρμογή των αντιμέτρων για την αποτροπή των εν λόγω απειλών, προτού αυτές, επηρεάσουν αρνητικά την εταιρεία. Η μοντελοποίηση αυτή, μπορεί να γίνει από τη σκοπιά ενός επιτιθέμενου, συλλέγοντας πληροφορίες σχετικά με τις μεθόδους που χρησιμοποιούν για τις ενέργειες εναντίων διαφορετικών εταιρειών, ώστε στη συνέχεια να σχεδιαστούν αντίμετρα για τις μεθόδους αυτές. Επίσης, μπορεί να γίνει από την οπτική γωνία των περιουσιακών στοιχείων

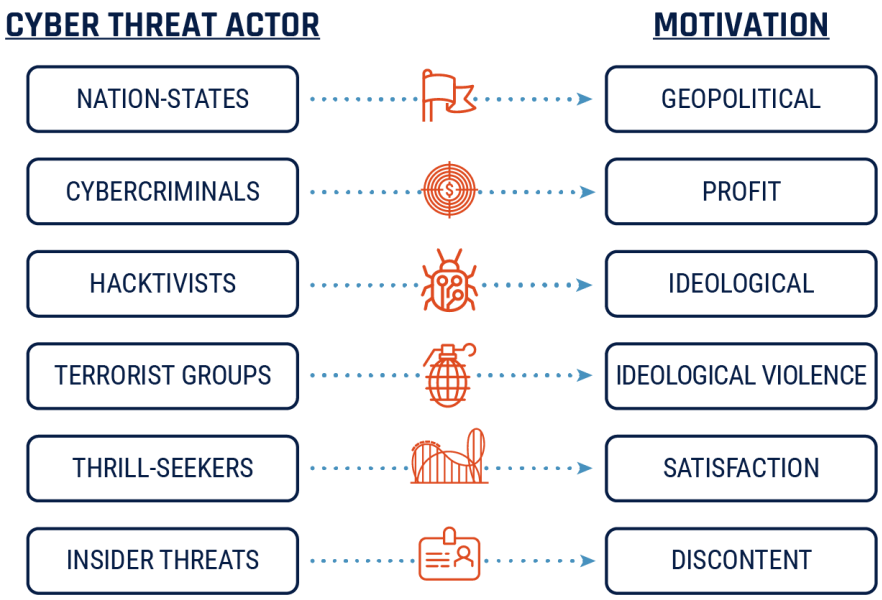
της εταιρείας, οπότε, μετά τον προσδιορισμό της σημαντικότητας αυτών, να εξεταστούν όλοι οι τρόποι που μπορεί να υπάρχουν για να παραβιαστούν αυτά και στη συνέχεια, να σχεδιαστούν οι κατάλληλοι έλεγχοι ασφαλείας για την αποτροπή των επιθέσεων.

Οι βασικές έννοιες για την κατανόηση ενός μοντέλου απειλών αφορούν τα εξής:

- Κυβερνοαπειλή
- Παράγοντες κυβερνοαπειλών
- Κίνητρα κυβερνοαπειλών
- Δραστηριότητες κυβερνοαπειλών

Γενικότερα μιλώντας, μία απειλή στον κυβερνοχώρο, νοείται ως μία δραστηριότητα η οποία αποσκοπεί στο να θέσει σε κίνδυνο την ασφάλεια ενός συστήματος πληροφοριών, αλλάζοντας τη διαθεσιμότητα, την ακεραιότητα ή το απόρρητο ενός συστήματος ή των πληροφοριών που περιέχει. Το περιβάλλον απειλής στον κυβερνοχώρο περιλαμβάνει τον διαδικτυακό χώρο όπου οι φορείς απειλών, μπορούν να ασκήσουν την κακόβουλη δραστηριότητα. Οι φορείς κυβερνοαπειλών μπορεί να είναι κράτη, ομάδες ή άτομα που, με κακόβουλη πρόθεση, στοχεύουν να εκμεταλλευτούν τα τρωτά σημεία, τη χαμηλή ευαισθητοποίηση στα ζητήματα κυβερνοασφάλειας ή τις τεχνολογικές εξελίξεις ώστε να αποκτήσουν μη εξουσιοδοτημένη πρόσβαση ή να επηρεάσουν με άλλο τρόπο τα δεδομένα, τις συσκευές και τα συστήματα των θυμάτων τους. Ο παγκοσμιοποιημένος χαρακτήρας του διαδικτύου, επιτρέπει σε αυτούς τους φορείς απειλών, να βρίσκονται φυσικά οπουδήποτε στον κόσμο και να εξακολουθούν να επηρεάζουν την ασφάλεια των πληροφοριακών συστημάτων, οπουδήποτε στον πλανήτη.

Οι φορείς των απειλών, μπορούν να κατηγοριοποιηθούν με βάση τα κίνητρά τους και, σε ορισμένο βαθμό, από την πολυπλοκότητά τους. Οι φορείς αυτοί, επιθυμούν την πρόσβαση σε συσκευές, επεξεργαστική ισχύ, υπολογιστικούς πόρους και πληροφορίες για διαφορετικούς λόγους ο καθένας. Γενικά, κάθε τύπος φορέων απειλής στον κυβερνοχώρο έχει ένα συγκεκριμένο, πρωταρχικό κίνητρο και οι περισσότεροι, παρουσιάζονται στην Εικόνα 4.



Εικόνα 4: Φορείς απειλών και κίνητρα

Οι φορείς αυτοί, δεν είναι ίσιοι ως προς την ικανότητα και την πολυπλοκότητα και διαθέτουν μία σειρά πόρων, εκπαίδευσης και υποστήριξης για τις δραστηριότητές τους. Οι φορείς κυβερνοαπειλών, μπορεί να λειτουργούν μόνοι τους ή ως μέρος ενός μεγαλύτερου οργανισμού όπως για παράδειγμα ένα πρόγραμμα πηλοφοριών ενός εθνικού κράτους ή μία ομάδα οργανωμένου εγκλήματος. Μερικές φορές, ακόμη και εξελιγμένοι φορείς, χρησιμοποιούν λιγότερο εξελιγμένα και άμεσα διαθέσιμα εργαλεία και τεχνικές, επειδή αυτά, μπορεί να είναι ακόμη αποτελεσματικά για μία δεδομένη απαίτηση. Οι φορείς των κρατών είναι συχνά οι πιο εξελιγμένοι παράγοντες απειλής, με αφοσιωμένους πόρους και προσωπικό, και με αρκετά εκτεταμένο σχεδιασμό και συντονισμό. Πολλές φορές, έχουν επιχειρησιακές σχέσεις με οντότητες του ιδιωτικού τομέα και οργανωμένους εγκληματίες.

Οι φορείς απειλών που βρίσκονται στην ανώτερη βαθμίδα πολυπλοκότητας και δεξιοτήτων, ικανοί να χρησιμοποιούν προηγμένες τεχνικές για τη διεξαγωγή σύνθετων και παρατεταμένων εκστρατειών για την επίτευξη των στρατηγικών τους στόχων, συχνά αποκαλούνται APT (Advanced Persistent Threats). Αυτός ο χαρακτηρισμός, προορίζεται συνήθως για έθνη-κράτη ή ορισμένες εξαιρετικά ικανές ομάδες του οργανωμένου εγκλήματος.

Χωρίς όλες να είναι απόλυτα επιβεβαιωμένες, οι πιο γνωστές ομάδες APT παρουσιάζονται στον Πίνακα 1.

Ονομασία APT	Χώρα Προέλευσης	Χώρες Στόχοι	Βασικό Εργαλείο
Lazarus Group	Βόρεια Κορέα	Νότια Κορέα & ΗΠΑ	Ransomware
Equation Group	ΗΠΑ	Ιράν, Συρία & Αφγανιστάν	Spyware
Fancy Bear	Ρωσία	ΗΠΑ & Γερμανία	Spear-Phishing
Machete	Νότια Αμερική	Βενεζουέλα, Κολομβία, Νικαράγουα & Εκουαδόρ	Phishing
Elfin	Ιράν	Σαουδική Αραβία & ΗΠΑ	Shamoon, Mimikatz, Spyware

Πίνακας 1: Advanced Persistent Threats ομάδες

Από την άλλη, οι χακτιβιστές, οι τρομοκρατικές ομάδες και εκείνοι οι οποίοι αναζητούν απλά την αδρεναλίνη, βρίσκονται στο χαμηλότερο επίπεδο πολυπλοκότητας, καθώς συχνά, βασίζονται σε ευρέως διαθέσιμα εργαλεία που απαιτούν λίγες τεχνικές δεξιότητες για την εφαρμογή τους. Οι ενέργειές τους, τις περισσότερες φορές, δεν έχουν μόνιμες επιπτώσεις στους στόχους τους, πέραν της φήμης τους.

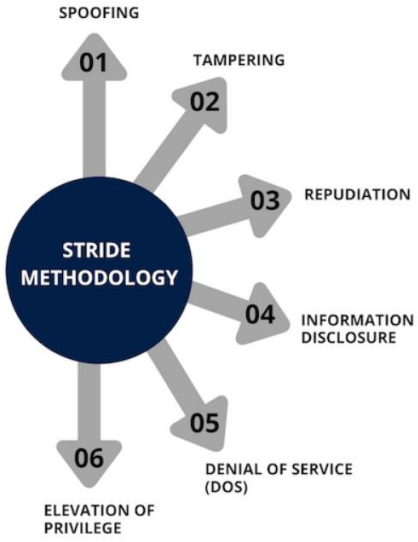
Οι εσωτερικές απειλές, προέρχονται από άτομα που εργάζονται στον εκάστοτε οργανισμό και είναι ιδιαίτερα επικίνδυνα, λόγω της πρόσβασής τους σε εσωτερικά δίκτυα που είναι προστατευμένα. Η πρόσβαση, αποτελεί βασικό στοιχείο για τους κακόβουλους παράγοντες απειλών και η πιθανή προνομιακή πρόσβαση, εξαλείφει την ανάγκη χρήσης άλλων απομακρυσμένων μέσων. Είναι πολύ συχνό, αυτού του είδους οι απειλές να προέρχονται από κάποιους δυσαρεστημένους υπαλλήλους.

Οι φορείς, διεξάγουν κακόβουλες δραστηριότητες εκμεταλλευόμενοι ευπάθειες ασφαλείας, χρησιμοποιώντας τεχνικές κοινωνικής μηχανικής ή χειραγωγώντας τα μέσα κοινωνικής δικτύωσης. Ένας πολύ ικανός κακόβουλος φορέας, θα επιλέξει πολύ προσεκτικά την τεχνική εκείνη η οποία είναι πιθανότερο να οδηγήσει σε μία επιτυχή εκμετάλλευση. Το συχνότερο πάντως είναι, μία μαζική και τυφλή προσπάθεια προς κάθε κατεύθυνση με την ελπίδα να εκμεταλλευτεί οποιοδήποτε μη ασφαλές δίκτυο ή βάση δεδομένων.

Σύμφωνα με άρθρο της Shevchenko στο Software Engineering Institute (2018) αυτά είναι τα πιο γνωστά μοντέλα απειλών.

- STRIDE
- MITRE ATT&CK
- PASTA
- TRIKE THREAT MODELING
- VAST
- DREAD
- OCTAVE MODEL

Το STRIDE είναι ένα μοντέλο που δημιουργήθηκε από την Microsoft και στοχεύει να βοηθήσει τις εφαρμογές να πληρούν τις οδηγίες ασφαλείας του CIA Triad. Εμπιστευτικότητα, ακεραιότητα και διαθεσιμότητα (Confidentiality – Integrity – Availability).



Εικόνα 5: STRIDE

Το MITRE ATT&CK είναι μία παγκόσμια βάση γνώσης τακτικών και τεχνικών των αντιπάλων. Χρησιμοποιείται στην ανάπτυξη μοντέλων απειλών και δίνει μία λεπτομερή περιγραφή των κοινών τεχνικών που χρησιμοποιούν οι άνθρωποι για να εισβάλλουν σε οργανισμούς.

Το PASTA προέρχεται από το Process for Attack Simulation και Threat Analysis και εστιάζει στην μοντελοποίηση απειλών από τη σκοπιά ενός χάκερ. Σκοπός του είναι να

παρέχει μια διαδικασία για την προσομοίωση επιθέσεων σε εφαρμογές, να αναλύει τις απειλές που προέρχονται από τις προσομοιώσεις αυτές και στη συνέχεια, να μετριάζει τον κίνδυνο που παρουσιάζουν αυτές οι επιθέσεις. Μέσω της διαδικασίας αυτής, επιτυγχάνεται ένα επαρκές επίπεδο ασφάλειας μέσω της επίλυσης των προβλημάτων που εντοπίστηκαν κατά την διάρκεια των προσημειώσεων αυτών.

Το μοντέλο Trike Threat είναι ένα ενοποιημένο εννοιολογικό πλαίσιο, για τον έλεγχο ασφαλείας και την άποψη της διαχείρισης κινδύνου μέσω της δημιουργίας μοντέλων απειλών με αξιόπιστο, επαναλαμβανόμενο τρόπο. Μια ομάδα ελέγχου ασφαλείας μπορεί αν το χρησιμοποιήσει για να περιγράψει πλήρως και με ακρίβεια τα χαρακτηριστικά ασφαλείας ενός συστήματος από την αρχιτεκτονική υψηλού επιπέδου έως τις λεπτομέρειες υλοποίησης χαμηλού επιπέδου. Επιτρέπει επίσης την επικοινωνία μεταξύ των μελών της ομάδας ασφαλείας και μεταξύ των ομάδων ασφαλείας και άλλων ενδιαφερόμενων μερών, παρέχοντας ένα συνεπές εννοιολογικό πλαίσιο.

Η μέθοδος μοντελοποίησης Visual, Agile και Simple Threat (VAST) βασίζεται στο ThreatModeler, μία αυτοματοποιημένη πλατφόρμα μοντελοποίησης απειλών. Η επεκτασιμότητα και η χρηστικότητά του, επιτρέπουν την υιοθέτησή του σε μεγάλους οργανισμούς σε ολόκληρη την υποδομή τους, για να παράγει αποτελεσματικά και αξιόπιστα αποτελέσματα για διαφορετικούς ενδιαφερόμενους φορείς. Αναγνωρίζοντας τις διαφορές στις λειτουργίες και τις ανησυχίες μεταξύ των ομάδων ανάπτυξης και υποδομής, το VAST, απαιτεί τη δημιουργία δύο τύπων μοντέλων. Μοντέλα απειλών εφαρμογών που χρησιμοποιούν διαγράμματα ροής διεργασιών αντιπροσωπεύοντας την αρχιτεκτονική άποψη και επίσης, μοντέλα επιχειρησιακών απειλών που δημιουργούνται από την σκοπιά του επιτιθέμενους με βάση τα DFD (????). Αυτή η συγκεκριμένη προσέγγιση, επιτρέπει την ενσωμάτωση του VAST και στον κύκλο ζωής της ανάπτυξης και του DevOps του οργανισμού.

Σχήμα 3: ThreatModeler

Το DREAD είναι ένα πλαίσιο μοντελοποίησης απειλών το οποίο δημιουργήθηκε από την Microsoft το 2003. Αν και η Microsoft έκτοτε το έχει εγκαταλείψει, επικαλούμενη ανησυχίες για την υποκειμενικότητά του, εξακολουθεί να χρησιμοποιείται σήμερα από μικρές επιχειρήσεις, εταιρείες του Fortune 500 αλλά και από τον στρατό. Το μοντέλο αυτό, αξιολογεί ποσοτικά τη σοβαρότητα μιας κυβερνοαπειλής, χρησιμοποιώντας ένα κλιμακωτό σύστημα αξιολόγησης το οποίο εκχωρεί αριθμητικές τιμές σε κατηγορίες κινδύνου. Το μοντέλο αυτό αποτελείται από πέντε κατηγορίες.

- Damage Potential
- Reproducibility
- Exploitability
- Affected Users
- Discoverability

Οι κατηγορίες αυτές αφορούν την πιθανή ζημιά που μπορεί να προκαλέσει μία συγκεκριμένη απειλή, το πόσο εύκολη είναι η αναπαραγωγιμότητα μιας επίθεσης, την ανάλυση των τρωτών σημείων του συστήματος για να εξακριβωθεί η ευαισθησία σε κυβερνοεπιθέσεις, το σύνολο των χρηστών που θα επηρεαστούν από μία τυχόν κυβερνοεπίθεση και τον προσδιορισμό της ευκολίας εντοπισμού, τυχόν ευάλωτων σημείων στην υποδομή του συστήματος.

Το Operationally Critical Threat, Asset και Vulnerability Evaluation (OCTAVE) είναι ένα πλαίσιο εντοπισμού και διαχείρισης κινδύνων ασφάλειας πληροφοριών. Καθορίζει μία ολοκληρωμένη μέθοδο αξιολόγησης, επιτρέποντας σε έναν οργανισμό τον προσδιορισμό των στοιχείων εκείνων που είναι σημαντικά για την αποστολή του, τις ενδεχόμενες απειλές για τα στοιχεία αυτά και τα τρωτά σημεία αυτών που μπορεί να τα εκθέσουν σε ορισμένες απειλές. Συνδυάζοντας όλα αυτά ο οργανισμός, αρχίζει να κατανοεί τις ευαίσθητες

πληροφορίες οι οποίες κινδυνεύουν και έτσι, μπορεί να σχεδιάσει και αν εφαρμόσει μια δεδομένη στρατηγικής προστασίας για την μείωση της συνολικής τους έκθεσης σε κίνδυνο.

1.4.1 Stride

Η μοντελοποίηση απειλών STRIDE είναι ένα συγκεκριμένο είδος μεθοδολογίας μοντελοποίησης απειλών. Πρόκειται για ένα "μνημονικό" έξι τύπων απειλών ασφαλείας. Κάθε γράμμα του STRIDE, όπως φαίνεται και στην Εικόνα 5, αντιπροσωπεύει έναν από τους έξι τύπους απειλών ασφαλείας.

- Spoofing
- Tampering
- Repudiation
- Information disclosure
- Denial of Service
- Elevation of Privilege

Η μοντελοποίηση απειλών STRIDE είναι χρήσιμη επειδή μπορεί να προβλέψει το τι μπορεί να πάει στραβά στην εφαρμογή, το σύστημα και ή την επιχειρησιακή διαδικασία που μοντελοποιούμε. Είναι μία απλή μεθοδολογία μοντελοποίησης απειλών, που μπορεί να χρησιμοποιηθεί από ένα ευρύ φάσμα ανθρώπων, συμπεριλαμβανομένων εκείνων που δεν έχουν παραδοσιακό υπόβαθρο στην ασφάλεια πληροφοριών. Η απλότητά του σημαίνει ότι μέλη διαφορετικών ομάδων μπορούν να εργαστούν με αυτό. Η ύπαρξη αυτών των διαφορετικών τύπου μελών, ακόμα και εκείνων χωρίς τεχνικό υπόβαθρο, βελτιώνει στην πραγματικότητα την συνολική ποιότητα και τα αποτελέσματα της μοντελοποίησης.

Το spoofing για παράδειγμα, συνίσταται στο να προσποιείσαι ότι είσαι κάποιος ή κάτι άλλο. Ένας επιτιθέμενος, θα χρησιμοποιήσει την ικανότητα αυτή για να εκτελέσει κακόβουλες ενέργειες για τις οποίες δεν θα έπρεπε να είναι ικανός ή θα την χρησιμοποιήσει επίσης, ως εφαλτήριο για περαιτέρω επιθέσεις. Παραδείγματα του spoofing, αποτελεί η αποστολή ηλεκτρονικού ταχυδρομείου ως άλλο πρόσωπο ή χρήστης, η επαναδημιουργία συναλλαγών ή η τροποποίηση λεπτομερειών ορισμένων συναλλαγών., η υποκλοπή ονόματος χρήστη και κωδικού πρόσβασης στο δίκτυο και επίσης, μία ψεύτικη φόρμα σύνδεσης στο διαδίκτυο για την συλλογή ονομάτων χρήστη και κωδικών πρόσβασης.

Το tampering συνίσταται στην αλλοίωση ή τροποποίηση δεδομένων, όταν αυτό δεν θα έπρεπε να είναι δυνατό. Κάτι τέτοιο, επηρεάζει την ακεραιότητα των δεδομένων. Παραδείγματα τέτοιων αλλοιώσεων μπορούν να οδηγήσουν σε άμεση πρόσβασης σε μια βάση δεδομένων μέσω μιας διεπαφής διαχείρισης, την τροποποίηση δεδομένων διαφορετικών χρηστών σε συγκεκριμένες εφαρμογές, στην αλλοίωση της επιχειρησιακής διαδικασίας εγγραφής του πελάτη και σε πιθανή αλλοίωση λεπτομερειών πληρωμής.

Το repudiation συνίσταται στη δυνατότητα να αρνηθεί κάποιος, όπως για παράδειγμα ένας χρήστης σε μια εφαρμογή ή ένας υπάλληλος σε μια επιχειρησιακή διαδικασία, αν μία ενέργεια πραγματοποιήθηκε ή όχι. Παραδείγματα τέτοιων ενεργειών αποτελούν η μη δυνατότητα επαλήθευσης δημιουργίας συναλλαγών σε εφαρμογή mobile banking και η αδυναμία επαλήθευσης των αποστολέων των μηνυμάτων σε μία εφαρμογή κοινωνικής δικτύωσης Business to Business (B2B).

Το Information Disclosure συνίσταται στην απόκτηση πρόσβασης σε εμπιστευτικές πληροφορίες κάτι το οποίο δεν θα έπρεπε να είναι δυνατό. Παραδείγματα αποκάλυψης ευαίσθητων πληροφοριών είναι σε πιθανές εφαρμογές τοπικής αυτοδιοίκησης για την προβολή εμπιστευτικών πληροφοριών σχετικά με ιδιώτες, σε πιθανή απόκτηση πρόσβασης διαχειριστή.

1.5 MITRE ATT&CK

Το MITRE ATT&CK αποτελεί μία επιμελημένη βάση γνώσης και ενός μοντέλου για την συμπεριφορά των επιτιθέμενων στον τομέα του κυβερνοχώρου. Σύμφωνα με τους Storm *et. al.* (2020 pp1-4) το μοντέλο αυτό αντικατοπτρίζει τις διάφορες φάσεις του κύκλου ζωής των επιθέσεων και τις πλατφόρμες που οι επιθέσεις αυτές, είναι γνωστό ότι στοχεύουν. Επικεντρώνεται στον τρόπο με τον οποίο οι επιτιθέμενοι θέτουν σε κίνδυνο και λειτουργούν εντός των πληροφοριακών συστημάτων. Η βάση αυτή προέκυψε από ένα αρχικό έργο το οποίο αφορούσε την τεκμηρίωση και κατηγοριοποίηση των τακτικών και των τεχνικών που χρησιμοποιούν οι επιτιθέμενοι στην στόχευση των λειτουργικών συστημάτων Microsoft Windows με σκοπό την βελτίωση ανίχνευσης κακόβουλων συμπεριφορών. Έκτοτε, έχει επεκταθεί και συμπεριλαμβάνει λειτουργικά συστήματα Linux και MacOS καθώς και κινητές συσκευές, συστημάτων τα οποία βασίζονται σε υπολογιστικό νέφος και τέλος, συστήματα βιομηχανικού ελέγχου.

Εικόνα 6: Mitre Att&ck

Σε υψηλό επίπεδο το ATT&CK είναι ένα μοντέλο συμπεριφοράς που αποτελείται από τα ακόλουθα βασικά στοιχεία.

- Τακτικές, που δηλώνουν τους βραχυπρόθεσμους, τακτικούς στόχους του επιτιθέμενου κατά την διάρκεια των επιθέσεων.
- Τεχνικές, που περιγράφουν τα μέσα με τα οποία οι αντίπαλοι επιτυγχάνουν τους τακτικούς στόχους.
- Υποτεχνικές, που περιγράφουν με περισσότερη ακρίβεια τα μέσα με τα οποία οι επιτιθέμενοι επιτυγχάνουν τους τακτικούς τους στόχους.
- Μία τεκμηριωμένη χρήση των τεχνικών από τους αντιπάλους, των διαδικασιών τους καθώς και άλλα μεταδεδομένα.

Το ATT&CK είναι οργανωμένο σε μια σειρά "τομέων τεχνολογίας" στο οικοσύστημα του οποίου λειτουργεί ένας επιτιθέμενος και παρέχει το σύνολο εκείνων των περιορισμών που χρειάζεται να παρακάμψει ή να εκμεταλλευτεί ο αντίπαλος για να επιτύχει ένα σύνολο στόχων. Οι τομείς αυτοί είναι τρεις και συγκεκριμένα το Enterprise, το Mobile και τα Industrial Control Systems (ICS).

Στο Enterprise περιλαμβάνονται οι ακόλουθες πλατφόρμες.

- Linux
- macOS
- Windows
- AWS
- Azure
- GCP
- SaaS
- Office365
- Azure AD

Και στο Mobile αντίστοιχα, περιλαμβάνονται τα Android και iOS.

1.5.1 Ιστορία του ATT&CK

Το ATT&CK δημιουργήθηκε από την ανάγκη της συστηματικής κατηγοριοποίησης της συμπεριφοράς των επιτιθέμενων στο πλαίσιο της διεξαγωγής δομημένων ασκήσεων προσομοίωσης των αντιπάλων σε ένα ερευνητικό περιβάλλον, το FMX. Το FMX ιδρύθηκε το 2010 και παρείχε τη δυνατότητα ενός "live εργαστηρίου" το οποίο επέτρεπε την πρόσβαση στους ερευνητές, σε έναν θύλακα παραγωγής του εταιρικού δικτύου του MITRE για την ανάπτυξη εργαλείων καθώς και για την δοκιμή και την τελειοποίηση ιδεών σχετικά με τον καλύτερο εντοπισμό απειλών. Η νοοτροπία που επιλέχθηκε να ακολουθηθεί για τον ταχύτερο εντοπισμό, προηγμένων και μόνιμων απειλών ήταν εκείνη της υπόθεσης, πως ένα σύστημα έχει ήδη παραβιαστεί.

1.5.2 Περιπτώσεις χρήσης του ATT&CK

Οι περιπτώσεις χρήσης του ATT&CK μπορούν να κατηγοριοποιηθούν ως εξής

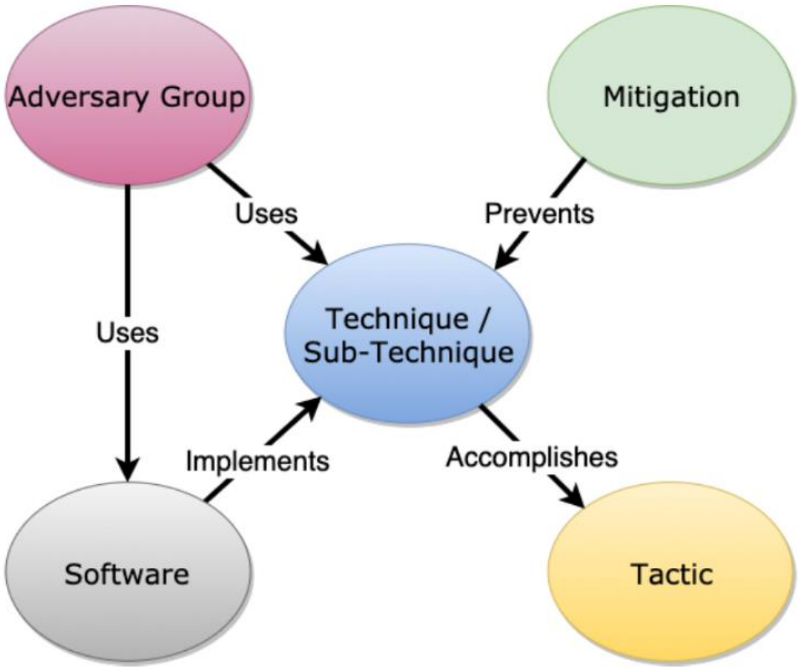
- Εξομοίωση αντιπάλου
- Red Teaming
- Ανάλυση συμπεριφοράς
- Αξιολόγηση δυνατοτήτων άμυνας
- Αξιολόγηση ωριμότητας ενός τμήματος SOC
- Εμπλουτισμός πληροφοριών που αφορούν κυβερνοαπειλές

Η εξομοίωση αντιπάλων, αφορά την δημιουργία σεναρίων εξομοίωσης για τον έλεγχο και την επαλήθευση της άμυνας, έναντι ορισμένων κοινών τεχνικών των αντιπάλων. Με τις δυνατότητες αυτές μπορούν να δημιουργηθούν ακόμα και προφίλ συγκεκριμένων ομάδων. Το Red teaming, αφορά την δημιουργία πλάνων και την οργάνωση επιθέσεων, για την αποφυγή ορισμένων αμυντικών μέτρων, που ενδέχεται να υπάρχουν σε ένα δίκτυο. Η ανάλυση συμπεριφοράς, αφορά τον εντοπισμό δυνητικά κακόβουλης δραστηριότητας μέσα σε ένα σύστημα, που ενδέχεται να μην βασίζεται σε προηγούμενη γνώση των εργαλείων ή των δεικτών ενός αντιπάλου. Έτσι, το ATT&CK χρησιμοποιείται ως εργαλείο για την ανάλυση συμπεριφοράς που δεν συνηθίζεται σε ένα περιβάλλον. Η αξιολόγηση δυνατοτήτων της άμυνας ενός οργανισμού, είναι χρήσιμη έτσι ώστε να υπάρξει επένδυση και βελτίωση ύστερα από τον εντοπισμό πιθανών κενών ασφαλείας στον οργανισμό αυτό.

Η αξιολόγηση ωριμότητας ενός τμήματος SOC, χρησιμοποιείται για να ανιχνεύσει, να κατανοήσει και να ανταποκρίνεται σε μεταβαλλόμενες απειλές του δικτύου με την πάροδο του χρόνου. Τέλος, όσον αφορά τον εμπλουτισμό των πληροφοριών το ATT&CK είναι χρήσιμο για την κατανόηση και την τεκμηρίωση των προφίλ των επιτιθέμενων ομάδων. Υπάρχει καλύτερη κατανόηση και χαρτογράφηση των κοινών συμπεριφορών των ομάδων αυτών με αποτέλεσμα, την αποτελεσματικότερη άμυνα σε περίπτωση επιθέσεων τους.

1.5.3 Μοντέλο και Matrix του ATT&CK

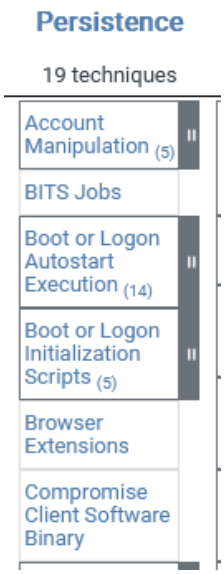
Βάση του MITRE ATT&CK, αποτελεί το σύνολο τεχνικών και επιμέρους τεχνικών που αντιπροσωπεύουν ενέργειες που μπορούν να εκτελέσουν οι αντίπαλοι για να επιτύχουν στόχους. Αυτοί οι στόχοι, αντιπροσωπεύονται από τις κατηγορίες τακτικής στις οποίες εμπίπτουν οι τεχνικές και οι επιμέρους τεχνικές. Αυτή η σχετικά απλή αναπαράσταση, επιτυγχάνει μία χρήσιμη ισορροπία μεταξύ επαρκών τεχνικών λεπτομερειών σε επίπεδο τεχνικής και του πλαισίου γύρω από το γιατί πραγματοποιούνται ενέργειες σε επίπεδο τακτικής.



Σχήμα 4: Σχέσεις του μοντέλου Att&ck

Η σχέση μεταξύ τακτικών, τεχνικών και επιμέρους τεχνικών απεικονίζεται μέσω ενός πίνακα, του ATT&CK Matrix. Για παράδειγμα όπως παρουσιάζεται και στην Εικόνα 7, όσον αφορά την τακτική Persistence δηλαδή τον στόχο του αντιπάλου να “παραμείνει” στον στόχο υπάρχουν ορισμένες από τις εξής τεχνικές:

- BITS Jobs
- Boot or Logon AutoStart Execution
- Browser Extensions
- Create Account
- PRE-OS Boot
- Scheduled Task/Job
- Office Application Startup
- Valid Accounts



Εικόνα 7: Persistence τεχνικές

Καθένα από αυτά, αποτελεί μία ενιαία τεχνική που μπορούν να χρησιμοποιήσουν οι επιτιθέμενοι έτσι ώστε να πετύχουν τον στόχο της "παραμονής". Επιπλέον, ορισμένες τεχνικές μπορούν να αναλυθούν και σε επιμέρους τεχνικές, περιγράφοντας με περισσότερες λεπτομέρειες το πως μπορούν να εκτελεστούν αυτές οι συμπεριφορές. Για παράδειγμα στο Scheduled Task/Job, όπως φαίνεται και στην Εικόνα 8, βλέπουμε υποκατηγορίες που αφορούν (At) Windows, (At) Linux, Systemd Timers, Cron κλπ.

Εικόνα 8: Scheduled Task/Job υποτεχνικές

Οι τακτικές, αντιπροσωπεύουν το “γιατί” μίας τεχνικής ή υπο-τεχνικής. Αποτελεί τον λόγο για την εκτέλεση της όποιας ενέργειας. Κάθε τεχνική, περιέχει έναν ορισμό που περιγράφει την κατηγορία και χρησιμεύει ως οδηγός για το ποιες τεχνικές πρέπει να περιλαμβάνονται στην τεχνική. Για παράδειγμα η εκτέλεση (Execution) ορίζεται ως μία τακτική που αντιπροσωπεύει υπο-τεχνικές οι οποίες έχουν ως αποτέλεσμα την εκτέλεση κώδικα ελεγχόμενου από τον αντίπαλο σε ένα τοπικό ή και απομακρυσμένο σύστημα. Αυτή η τακτική, χρησιμοποιείται συχνά σε συνδυασμό με την αρχική πρόσβαση (Initial Access) ως μέσο εκτέλεσης κώδικα, μόλις αποκτηθεί πρόσβαση στο σύστημα καθώς και με την επέκταση πρόσβασης (Lateral Movement) για την επέκταση πρόσβασης και σε υπόλοιπα, απομακρυσμένα συστήματα του δικτύου.

2. Μηχανική μάθηση

Σύμφωνα με τους Naqa et all (2015, pp.3-4), η μηχανική μάθηση είναι ένας εξελισσόμενος κλάδος υπολογιστικών αλγορίθμων, που έχουν σχεδιαστεί για να μιμούνται την ανθρώπινη νοημοσύνη μαθαίνοντας από το περιβάλλον. Αποτελεί ιδιαίτερα σημαντικό τομέα, ιδίως στην εποχή μας, των "big data". Τεχνικές που βασίζονται στη μηχανική μάθηση, εφαρμόζονται με επιτυχία σε πολλούς τομείς όπως η αναγνώριση προτύπων, η μηχανική όραση, η μηχανική διαστημικών σκαφών, ο κλάδος της οικονομίας, η ψυχαγωγία, η βιοιατρική, οι ιατρικές εφαρμογές κλπ.

Το πεδίο της μηχανικής μάθησης, έχει λάβει αρκετούς επίσημους ορισμούς στην διαθέσιμη βιβλιογραφία. Αυτοί είναι ορισμένοι από τους πιο γνωστούς ορισμούς της.

- Ένα πεδίο μελέτης που δίνει στους υπολογιστές τη δυνατότητα να μαθαίνουν, χωρίς να είναι ρητά προγραμματισμένοι. Mahesh (2020)
- Ένα πρόγραμμα υπολογιστή λέγεται ότι μαθαίνει από την εμπειρία (E) σε σχέση με κάποια κατηγορία εργασιών (T) και το μέτρο απόδοσης (P), εάν η απόδοσή του σε εργασίες στο T, όπως μετρούμενη με P, βελτιώνεται με την εμπειρία. Naqa et all (2015)

Όλοι οι από πάνω ορισμοί, μοιράζονται την έννοια της καθοδήγησης υπολογιστών για την πιο έξυπνη εκτέλεση εργασιών, πέρα από τις παραδοσιακές μεθόδους. Πρόκειται για έναν από τους ταχύτερα αναπτυσσόμενους τεχνικούς τομείς και βρίσκεται στην διασταύρωση της επιστήμης των υπολογιστών, της στατιστικής, της τεχνητής νοημοσύνης και της επιστήμης των δεδομένων. Η ανάπτυξη στην μηχανική μάθηση οφείλεται τόσο στην δημιουργία νέων αλγορίθμων μάθησης και θεωρίας, όσιο και στην συνεχιζόμενη έκρηξη της διαθεσιμότητας online δεδομένων και χαμηλού κόστους υπολογισμών τους. Η υιοθέτηση μεθόδων μηχανικής μάθησης, συναντάται σε ολόκληρη την επιστήμη, την τεχνολογία και το εμπόριο, οδηγώντας σε πιο τεκμηριωμένη λήψη αποφάσεων σε πολλούς τομείς της ζωής, όπως η υγειονομική περίθαλψη, η μεταποίηση, η εκπαίδευση, η χρηματοοικονομική μοντελοποίηση, η αστυνόμευση, το μάρκετινγκ κ.α.

Ή έναρξη της μηχανικής μάθησης μπορεί να εντοπιστεί στον 17^{ο} αιώνα και στην ανάπτυξη μηχανών που μπορούν να μιμηθούν την ανθρώπινη ικανότητα να προσθέτουν και να αφαιρούν, κάτι που συνέβη από τον Pascal και Leibniz. Στη σύγχρονη ιστορία, ο Arthur Samuel (1957) της IBM, επινόησε τον όρο μηχανική μάθηση και απέδειξε ότι οι υπολογιστές θα μπορούσαν να προγραμματιστούν για να παίζουν το γνωστό παιχνίδι

“checkers”. Ακολουθήθηκε από την ανάπτυξη του perceptron από τον Rosenblatt (1958) και αποτελούσε μία από τις πρώτες αρχιτεκτονικές νευρωνικών δικτύων το 1958. Μία επίσης σημαντική ανακάλυψη επιτεύχθηκε το 1975 με την ανάπτυξη του πολυστρωματικού perceptron από τον Werbos (1974) και ακολούθησε η ανάπτυξη δέντρων απόφασης από τον Quinlan (1986) και μηχανών υποστήριξης διανυσμάτων από τους Cortes et all (1995). Ακολούθησαν οι προτάσεις μηχανικής μάθησης Adaboost από Schapire (1999) και Random Forests από Breiman (2001). Τέλος, πρόσφατα, από Hinton (2007) έκαναν την εμφάνισή τους κατανεμημένοι πολυεπίπεδοι αλγόριθμοι “βαθιάς” μάθησης (deep learning).

Η μηχανική μάθηση, σύμφωνα με τον Jordan (2015) επικεντρώνεται σε δύο αλληλένδετα ερωτήματα. Το πρώτο ερώτημα είναι το πως μπορεί να κατασκευάσει κάποιος συστήματα υπολογιστών που βελτιώνονται αυτόματα μέσω της εμπειρίας και ποιοι είναι οι θεμελιώδεις στατιστικοί υπολογιστικοί θεωρητικοί νόμοι πληροφορίας, που διέπουν όλα τα συστήματα μάθησης, συμπεριλαμβανομένων των υπολογιστών, των αθρόων και των οργανισμών. Η μελέτη της μηχανικής μάθησης είναι σημαντική, τόσο για την αντιμετώπιση αυτών των θεμελιωδών επιστημονικών και μηχανικών ερωτημάτων, όσο και για το εξαιρετικά πρακτικό λογισμικό υπολογιστών που έχει παραχθεί και εφαρμοστεί σε πολλές εφαρμογές.

Έχει σημειώσει δραματική πρόοδο τις τελευταίες δύο δεκαετίες, από εργαστηριακή περιέργεια σε μία πρακτική τεχνολογία με ευρεία εμπορική χρήση. Στο πλαίσιο της τεχνητής νοημοσύνης η μηχανική μάθηση έχει αναδειχθεί ως η μέθοδος επιλογής για την ανάπτυξη πρακτικού λογισμικού για την όραση υπολογιστών, την αναγνώρισης ομιλίας, την επεξεργασία φυσικής γλώσσας, τον έλεγχο ρομπότ καθώς και δεκάδες άλλες εφαρμογές. Πολλοί προγραμματιστές συστημάτων τεχνητής νοημοσύνης αναγνωρίζουν τώρα ότι, για πολλές εφαρμογές, μπορεί να εκπαιδευτεί ένα σύστημα, δείχνοντάς του παραδείγματα της επιθυμητής συμπεριφοράς εισόδου-εξόδου από το να το προγραμματίσει χειροκίνητα, προβλέποντας την επιθυμητή απόκριση για όλες τις πιθανές εισόδους. Η επίδραση της μηχανικής μάθησης έχει επίσης γίνει ευρέως αισθητή, σε όλη την επιστήμη των υπολογιστών, σε μία σειρά από βιομηχανίες, που οι λειτουργίες τους βασίζονται και επηρεάζονται από την ύπαρξη δεδομένων. Αντίστοιχα, υπήρξε ένα ευρύ φάσμα επιπτώσεων στις εμπειρικές επιστήμες, από την βιολογία έως την κοσμολογία και τις κοινωνικές επιστήμες, καθώς αναπτύχθηκαν μέθοδοι μηχανικής μάθησης για την ανάλυση πειραματικών δεδομένων υψηλής απόδοσης με νέους τρόπους.

Το πρόβλημα μάθησης μπορεί να οριστεί ως το πρόβλημα της βελτίωσης κάποιου μέτρου απόδοσης κατά την εκτέλεση κάποιας εργασίας, μέσω κάποιου είδους εκπαιδευτικής εμπειρίας. Ως παράδειγμα, στην εκμάθηση ανίχνευσης απάτης πιστωτικών καρτών, στόχος είναι η κατηγοριοποίηση ετικέτας "απάτης" ή "όχι απάτης" σε κάθε συναλλαγή κάρτας. Η εκμάθηση αυτή πραγματοποιείται, χρησιμοποιώντας μία συλλογή ιστορικών συναλλαγών των πιστωτικών καρτών κάθε μία από τις οποίες έχει ήδη χαρακτηριστεί ως απάτη ή όχι.

Ένας αλγόριθμος μηχανικής μάθησης είναι μια υπολογιστική διαδικασία που χρησιμοποιεί δεδομένα εισόδου για να επιτύχει μια επιθυμητή εργασία, χωρίς να είναι κυριολεκτικά (hard coded) προγραμματισμένος για την επιθυμητή αυτή εργασία. Αυτοί οι αλγόριθμοι θεωρούνται (soft coded) καθώς μεταβάλλουν ή προσαρμόζουν αυτόματα την αρχιτεκτονική τους μέσω της επανάληψης, δηλαδή της εμπειρίας. Με τον τρόπο αυτό, γίνονται όλο και καλύτεροι στην επίτευξη της επιθυμητής εργασίας. Η διαδικασία προσαρμογής ονομάζεται εκπαίδευση, στην οποία παρέχονται δείγματα δεδομένων εισόδου μαζί με τα επιθυμητά τους αποτελέσματα. Στην συνέχεια, ο αλγόριθμος διαμορφώνεται βέλτιστα έτσι ώστε, όχι μόνο να παράγει το επιθυμητό αποτέλεσμα όταν παρουσιάζεται με τις εισαγωγές εκπαίδευσης αλλά να μπορεί να γενικεύει έτσι ώστε να παράγει το επιθυμητό αποτέλεσμα και από νέα δεδομένα, που προηγουμένως δεν είχαν εμφανιστεί. Αυτή η εκπαίδευση αποτελεί το "μαθησιακό" μέρος της μηχανικής μάθησης. Η εκπαίδευση, δεν χρειάζεται αν περιορίζεται σε μια αρχική προσαρμογή, σε ένα πεπερασμένο διάστημα. Όπως και με τους ανθρώπους, ένας καλό αλγόριθμος μπορεί να εξασκήσει τη "δια βίου" μάθηση, καθώς αυτός επεξεργάζεται νέα δεδομένα και μαθαίνοντας από τα λάθη του.

Υπάρχουν πολλοί τρόποι με τους οποίους ένας υπολογιστικός αλγόριθμος, μπορεί να προσαρμοστεί στην εκπαίδευση. Τα δεδομένα εισόδου, μπορούν να επιλεγούν και να σταθμιστούν για να παρέχουν τα πιο καθοριστικά αποτελέσματα. Ο αλγόριθμος επίσης, μπορεί να έχει μεταβλητές αριθμητικές παραμέτρους που προσαρμόζονται μέσω επαναληπτικής βελτιστοποίησης. Επιπλέον, μπορεί να έχει ένα δίκτυο πιθανών υπολογιστικών διαδρόμων, που οργανώνονται για βέλτιστα αποτελέσματα. Τέλος, μπορεί να προσδιορίσει τις κατανομές πιθανοτήτων από τα δεδομένα εισόδου και να τις χρησιμοποιήσει για να προβλέψει τα αποτελέσματα.

2.1 Τύποι μηχανικής μάθησης

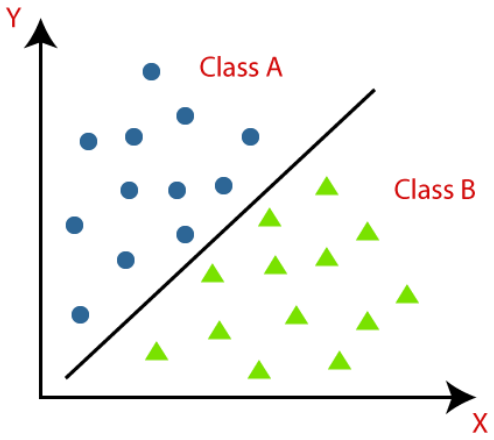
Σύμφωνα με τον Murphy (2012, pp.33-34) συνήθως, δύο είναι οι βασικοί τύποι στους οποίους χωρίζεται η μηχανική μάθηση. Στην προγνωστική ή εποπτευόμενη μάθηση, ο στόχος είναι να μάθουμε μία αντιστοίχιση από τις εισόδους x στις εξόδους y, δεδομένου επισημασμένου συνόλου ζευγών εισόδου-εξόδου.

$$D = \{(x_i, y_i)_{i=1}^N$$

Στην περίπτωση αυτή, το D αποτελεί το σύνολο της εκπαίδευσης και το N αποτελεί τον αριθμό των παραδειγμάτων της εκπαίδευσης. Στην απλούστερη μορφή, κάθε είδος εισόδου x_i είναι ένα διάνυσμα D-διάστατων αριθμών, όπως για παράδειγμα το μήκος και το πλάτος ενός φύλλου φυτού. Αυτά, ονομάζονται χαρακτηριστικά ή μεταβλητές. Ωστόσο, το x_i θα μπορούσε να είναι αρκετά πολύπλοκο αντικείμενο, όπως μία εικόνα, μία πρόταση, ένα μοριακό σχήμα, ένα γράφημα κλπ. Η μορφή της μεταβλητής εξόδου ή απόκρισης μπορεί να είναι οτιδήποτε, οι περισσότεροι μέθοδοι όμως υιοθετούν ότι το y_i είναι μία κατηγορική μεταβλητή από ένα πεπερασμένο σύνολο

$$y_i \in \{1, \ldots, C\}$$

όπως γάτα ή σκύλος ή ότι το y_i είναι μία βαθμίδα πραγματικής αξίας, όπως για παράδειγμα το επίπεδο αξίας ενός ακινήτου. Όταν το y_i είναι κατηγορικό το πρόβλημα θεωρείται ταξινόμησης ή αναγνώρισης προτύπων και όταν το y_i είναι πραγματική τιμή, το πρόβλημα είναι γνωστό ως παλινδρόμηση. Υπάρχει μία ακόμη παραλλαγή, γνωστή ως παλινδρόμηση κατά σειρά όπου σε αυτή, ο χώρος της ετικέτας Y έχει μία φυσική σειρά όπως για παράδειγμα οι βαθμοί από A-F. Με το C να αποτελεί τον αριθμό των κλάσεων, εάν είναι 2 τότε αναφερόμαστε σε δυαδική ταξινόμηση και εάν C > 2 τότε αναφερόμαστε σε ταξινόμηση πολλών κλάσεων. Εάν οι ετικέτες κλάσεων δεν είναι αμοιβαία αποκλειόμενες, δηλαδή μία εικόνα να περιέχει και έναν σκύλο και μία γάτα, τότε αναφερόμαστε σε ταξινόμηση πολλαπλών ετικετών.



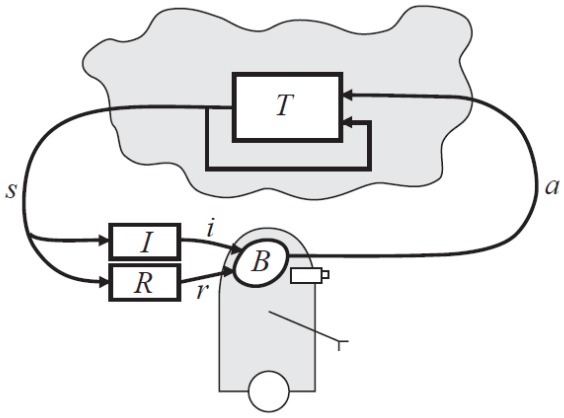
Σχήμα 5: Αλγόριθμος classification

Ο δεύτερος, κύριος τύπος μηχανικής μάθησης είναι η περιγραφική ή χωρίς επίβλεψη προσέγγιση μάθησης. Στην περίπτωση αυτή, δίνονται μόνο οι είσοδοι

$$D = \{x_i\}_{i=1}^{N}$$

και ο στόχος είναι η εύρεση μοτίβων που παρουσιάζουν ενδιαφέρον, στα εισαγόμενα αυτά δεδομένα. Αυτό, μερικές φορές ονομάζεται ανακάλυψη γνώσης. Το πρόβλημα αυτό, δεν είναι ένα καλά καθορισμένο πρόβλημα καθώς δεν γνωρίζουμε τη μοτίβα αναζητούνται και δεν υπάρχει κάποια προφανής μέτρηση σφάλματος ώστε να χρησιμοποιηθεί. Η εργασία μπορεί να επισημοποιηθεί ως μία εκτίμηση πυκνότητας επιδιώκοντας την δημιουργία μοντέλου στη μορφή $p(x_i|\theta)$ σημαίνοντας μία άνευ όρων εκτίμηση της πυκνότητας. Το x_i αποτελεί ένα διάνυσμα χαρακτηριστικών, σημαίνοντας πως είναι απαραίτητη η δημιουργία μοντέλων πολυμεταβλητών πιθανοτήτων.

Ένας τρίτος τύπος μάθησης είναι η ενισχυτική μάθηση που περιλαμβάνει την εκμάθηση λήψης αποφάσεων μέσω της αλληλεπίδρασης με ένα περιβάλλον και την λήψη αμοιβών ή ποινών, με βάση τις ενέργειές της. Σύμφωνα με τους Kaelbling et. al. (1996) η ενισχυτική μάθηση είναι το πρόβλημα που αντιμετωπίζει ένας "agent" που πρέπει να μάθει τη συμπεριφορά μέσω αλληλεπιδράσεων δοκιμής και σφάλματος σε ένα δυναμικό περιβάλλον.



Σχήμα 6: Τυπικό μοντέλο ενισχυτικής μάθησης

Θεωρείται κυρίως ως μία κατηγορία προβλημάτων παρά ως ένα σύνολο τεχνικών. Υπάρχουν δύο κύριες στρατηγικές για την επίλυση προβλημάτων ενισχυτικής μάθησης. Η πρώτη είναι μία αναζήτηση στον χώρο των συμπεριφορών για να βρεθεί ποια έχει καλή απόδοση στο περιβάλλον, κάτι που έχει υιοθετηθεί σε εργασίες σε γενετικούς αλγορίθμους και γενετικό προγραμματισμό. Η δεύτερη είναι η χρήση στατιστικών τεχνικών και μεθόδων δυναμικού προγραμματισμού για την εκτίμηση της χρησιμότητας της ανάληψης δράσεων.

Σε ένα τυπικό μοντέλο ενισχυτικής μάθησης, ένας “agent” συνδέεται με το περιβάλλον του μέσω «αντίληψης» και «δράσης» όπως φαίνεται και στο Σχήμα 6. Σε κάθε βήμα αλληλεπίδρασης ο “agent” λαμβάνει ως είσοδο i, κάποια ένδειξη της τρέχουσας κατάστασης s, του περιβάλλοντος και στη συνέχεια, επιλέγει μία ενέργεια a για να παράξει έξοδο. Η ενέργεια, αλλάζει την κατάσταση περιβάλλοντος και η τιμή αυτή της μετάβασης κατάστασης, κοινοποιείται στον “agent” μέσω ενός σήματος r. Η επιλογή των ενεργειών, θα πρέπει να αυξάνει το μακροπρόθεσμο άθροισμα των τιμών και αυτό θα επιτευχθεί με την πάροδο του χρόνο, έπειτα από συστηματικές δοκιμές και σφάλματα.

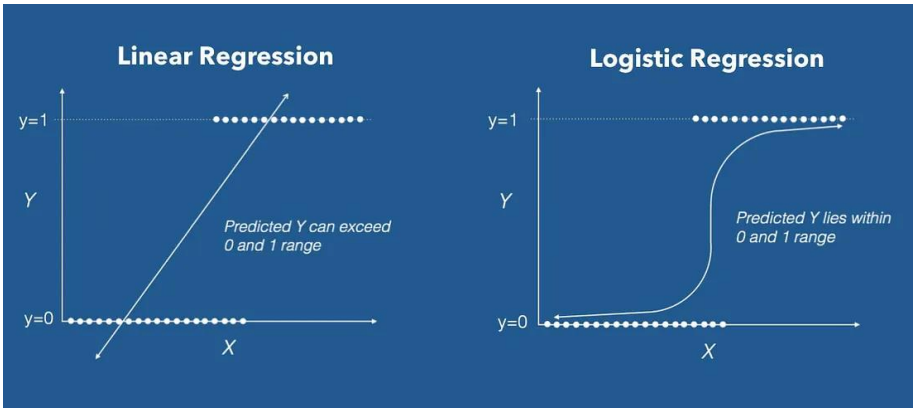
2.2 Αλγόριθμοι μηχανικής μάθησης

Η γραμμική παλινδρόμηση (Linear Regression) σύμφωνα με Maulud et al. (2020) είναι ένας ευρέως χρησιμοποιούμενος αλγόριθμος στη μηχανική μάθηση, για την πρόβλεψη μεταβλητών συνεχούς εξόδου και αφορά την εύρεση μια γραμμικής σχέσης μεταξύ ενός ή περισσότερων προγνωστικών. Λειτουργεί προσαρμόζοντας μία γραμμική εξίσωση σε ένα σύνολο δεδομένων εισόδου, τα οποία στη συνέχεια μπορούν να χρησιμοποιηθούν για προσπάθεια πρόβλεψης σε νέα δεδομένα. Ο αλγόριθμος, επιδιώκει να ελαχιστοποιήσει το

άθροισμα των τετραγωνικών σφαλμάτων μεταξύ της προβλεπόμενης και της πραγματικές εξόδου για κάθε σημείο δεδομένων. Αυτό, επιτυγχάνεται μέσω μιας διαδικασίας που ονομάζεται παλινδρόμηση ελαχίστων τετραγώνων. Τα μοντέλα αυτά είναι ιδιαίτερα χρήσιμα, σε καταστάσεις όπου υπάρχει γραμμική σχέση μεταξύ των μεταβλητών εισόδου και εξόδου και όπου υπάρχει ανάγκη πρόβλεψης αριθμητικών τιμών. Χρησιμοποιούνται σε ένα ευρύ φάσμα εφαρμογών, όπως η πρόβλεψη πωλήσεων, η εκτίμηση της ζήτησης της αγοράς και η πρόβλεψη οικονομικών τάσεων.

$$y = \beta 0 + \beta 1x + \varepsilon$$

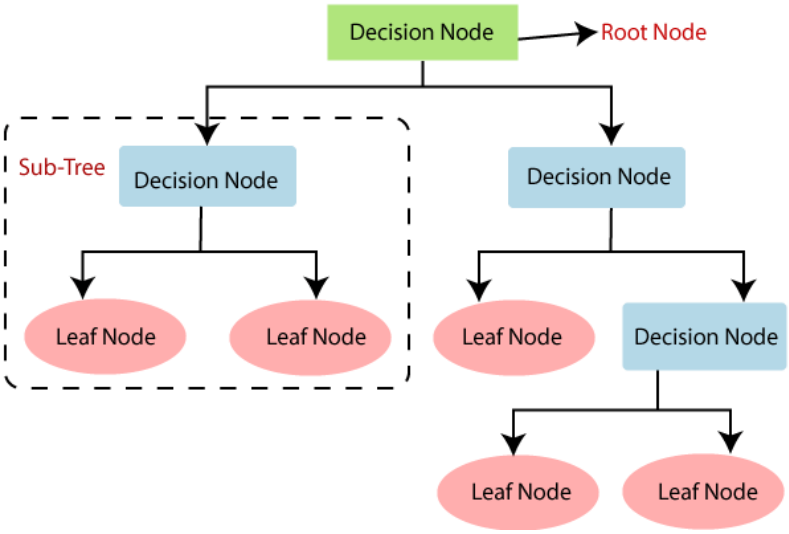
Η λογιστική παλινδρόμηση (Logistic Regression) σύμφωνα με τον Hosmer et. al. (2013) αποτελεί έναν αλγόριθμο ταξινόμησης, που χρησιμοποιείται για την πρόβλεψη της πιθανότητας μια δυαδικής ή διχοτομικής μεταβλητής αποτελέσματος, με βάση μία ή περισσότερες μεταβλητές πρόβλεψης. Λειτουργεί, προσαρμόζοντας μια λογιστική συνάρτηση στα δεδομένα η οποία αντιστοιχίζει τις μεταβλητές εισόδου σε μια τιμή πιθανότητας μεταξύ 0 και 1. Αποτελεί δημοφιλή αλγόριθμο σε πολλές εφαρμογές, όπως η ιατρική διάγνωση, η ανάλυση πιστωτικού κινδύνου, η τμηματοποίηση πελατών κ.α.



Σχήμα 7: Λογιστική και γραμμική παλινδρόμηση

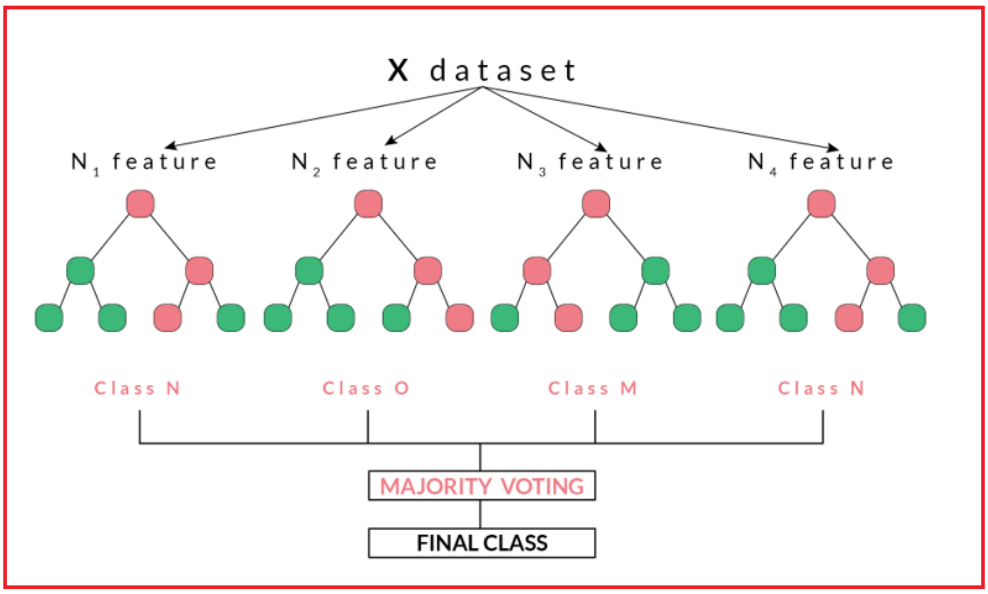
Τα δέντρα αποφάσεων (Decision Trees) είναι ένας ακόμη δημοφιλής αλγόριθμος μηχανικής μάθησης, για εργασίες ταξινόμησης αλλά και παλινδρόμησης. Σύμφωνα με τον Kotsiantis (2013) είναι διαδοχικά μοντέλα, τα οποία συνδυάζουν λογικά, μία ακολουθία απλών δοκιμών, με κάθε δοκιμή να συγκρίνει ένα αριθμητικό χαρακτηριστικό έναντι ενός συνόλου πιθανών τιμών. Τέτοιου είδους ταξινομητές έχουν το πλεονέκτημα ότι είναι πιο εύκολα κατανοητοί. Οι λογικοί κανόνες που ακολουθούνται από ένα δέντρο αποφάσεων, είναι πολύ πιο εύκολο να ερμηνευτούν από τα αριθμητικά βάρη για παράδειγμα των συνδέσεων μεταξύ

κόμβων ενός νευρωνικού δικτύου. Λειτουργούν με την αναδρομική κατανομή των δεδομένων σε υποσύνολα με βάση τις τιμές των χαρακτηριστικών εισόδου, με στόχο την ελαχιστοποίηση των προσμίξεων ή της αβεβαιότητας της μεταβλητής στόχου, σε κάθε υποσύνολο. Κάθε διαχωρισμός, δημιουργεί έναν νέο κόμβο στο δέντρο, με τον κόμβο-ρίζα να αντιπροσωπεύει ολόκληρο το σύνολο δεδομένων και τους κόμβους φύλλων, να αντιπροσωπεύουν την τελική ταξινόμηση ή τις προβλέψεις παλινδρόμησης.



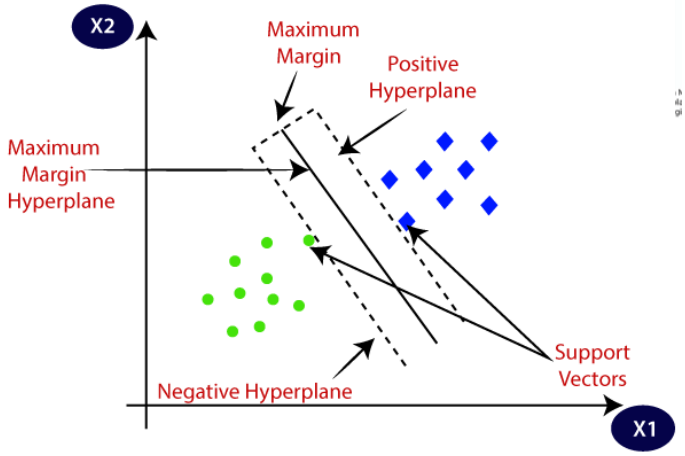
Σχήμα 8: Δέντρα αποφάσεων

Το Random Forest, σύμφωνα με Breiman (2001) είναι ένας αλγόριθμος μηχανικής μάθησης, που βασίζεται στον αλγόριθμο των δέντρων αποφάσεων. Δημιουργεί ένα σύνολο δέντρων αποφάσεων, όπου κάθε δέντρο εκπαιδεύεται σε ένα τυχαίο υποσύνολο των δεδομένων εισόδου και ένα τυχαίο υποσύνολο των χαρακτηριστικών εισόδου. Οι εσωτερικές εκτιμήσεις, παρακολουθούν το σφάλμα, την ισχύ και τη συσχέτιση και χρησιμοποιούνται για να δείξουν την απόκριση στην αύξηση του αριθμού των χαρακτηριστικών που χρησιμοποιούνται στη διαίρεση. Η τελική πρόβλεψη γίνεται με τη συγκέντρωση των προβλέψεων, όλων των μεμονωμένων δέντρων, είτε λαμβάνοντας την πλειοψηφία για εργασίες ταξινόμησης είτε τον μέσο όρο για εργασίες παλινδρόμησης. Έχει σχεδιαστεί για να μειώνει την υπερπροσαρμογή, να αυξάνει την ακρίβεια και τη σταθερότητα των προβλέψεων και επίσης, να χειρίζεται δεδομένα υψηλών διαστάσεων καθώς και εκείνα με υψηλό «θόρυβο».



Σχήμα 9: Random forest

Το Support Vector Machine (SVM) είναι ένας δημοφιλής αλγόριθμος μηχανικής μάθησης που κι αυτός, χρησιμοποιείται τόσο για εργασίες ταξινόμησης όσο και για εργασίες παλινδρόμησης. Σύμφωνα με Cortes (1995) Λειτουργεί βρίσκοντας ένα υπερεπίπεδο που μεγιστοποιεί το περιθώριο μεταξύ των δύο κλάσεων, όπου το περιθώριο ορίζεται ως η απόσταση μεταξύ του υπερεπίπεδου και των πλησιέστερων σημείων δεδομένων από κάθε κατηγορία. Μπορεί να χειριστεί τόσο γραμμικά όσο και μη γραμμικά διαχωρίσιμα δεδομένα, χρησιμοποιώντας συναρτήσεις πυρήνα, οι οποίες αντιστοιχίζουν τα χαρακτηριστικά εισόδου σε έναν χώρο υψηλότερης διάστασης όπου τα δεδομένα γίνονται γραμμικά διαχωρίσιμα. Είναι γνωστό για την ικανότητα να χειρίζεται δεδομένα υψηλών διαστάσεων καθώς και την ανθεκτικότητα σε ακραίες τιμές και «θόρυβο».

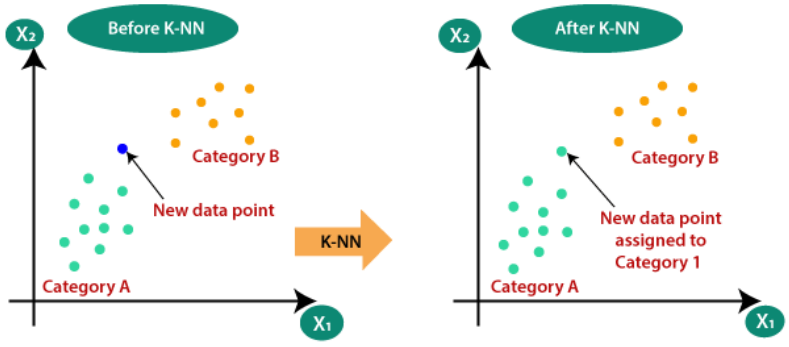


Σχήμα 10: Support vector machines

Σύμφωνα με τον Rish (2001) ο Naive Bayes, απλοποιεί σημαντικά την μάθηση, υποθέτοντας πως τα χαρακτηριστικά αποτελούν υπό όρους ανεξάρτητα μεταξύ τους κάτι που σημαίνει ότι η παρουσία ενός χαρακτηριστικού, δεν επηρεάζει την πιθανότητα των άλλων χαρακτηριστικών. Είναι ένα απλώς αλλά ισχυρός αλγόριθμος μηχανικής εκμάθησης που χρησιμοποιείται συνήθως για εργασίες ταξινόμησης κειμένου και φιλτραρίσματος ανεπιθύμητων μηνυμάτων. Βασίζεται στο θεώρημα του Bayes, το οποίο υπολογίζει την πιθανότητα μιας κλάσης δεδομένης ενός διανύσματος εισόδου.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

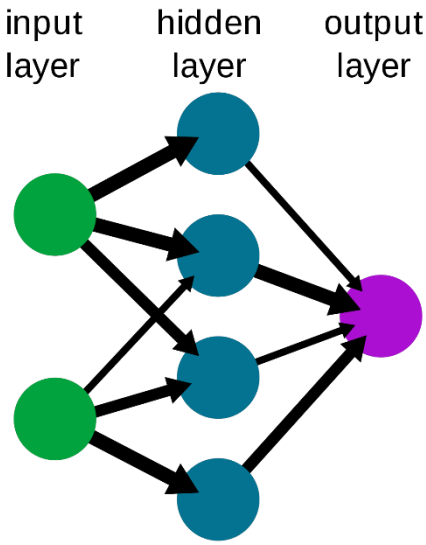
Σύμφωνα με τον Guo et. al. (2003) ο αλγόριθμος k-Nearest-Neighboors είναι μια παραμετρική μέθοδος ταξινόμησης η οποία είναι απλή αλλά αποτελεσματική σε πολλές περιπτώσεις. Για να ταξινομηθεί μία εγγραφή δεδομένων t ανακτώνται οι k πλησιέστεροι γείτονές της και έτσι, σχηματίζεται μία «γειτονιά» του t. Η απόφαση ταξινόμησης λαμβάνεται μέσω της πλειοψηφίας των εγγραφών δεδομένων στη γειτονιά με ή χωρίς να λαμβάνεται υπόψη η στάθμιση με βάση την απόσταση. Για την εφαρμογή του kNN σημαντικό ρόλο παίζει η κατάλληλη επιλογή της τιμής k.



Σχήμα 11: Αλγόριθμος kNN

Τα νευρωνικά δίκτυα σύμφωνα με τον Maind et. al. (2014) είναι ένας τύπος αλγόριθμου μηχανικής μάθησης που μοντελοποιεί τη συμπεριφορά του ανθρώπινου εγκεφάλου ώστε να μαθαίνει και να πετυχαίνει προβλέψεις από δεδομένα. Αποτελείται από πολλαπλά στρώματα διασυνδεδεμένων κόμβων ή νευρώνων, καθένας από τους οποίους επεξεργάζεται την είσοδο και την μεταβιβάζει στο επόμενο επίπεδο. Το νευρωνικό δίκτυο μαθαίνει

προσαρμόζοντας τα βάρη και το bias των νευρώνων μέσω μιας διαδικασίας που ονομάζεται backpropagation η οποία ελαχιστοποιεί το σφάλμα μεταξύ της προβλεπόμενης εξόδου και της πραγματικής εξόδου. Μπορούν να χειριστούν πολύπλοκες και μη γραμμικές σχέσεις στα δεδομένα και μπορούν να χρησιμοποιηθούν τόσο για εργασίες ταξινόμησης όσο και παλινδρόμησης.



Σχήμα 12: Νευρωνικό δίκτυο

2.3 Επεξεργασία φυσικής γλώσσας

Η επεξεργασία φυσικής γλώσσας σύμφωνα με την Hirschberg et. al. (2015) χρησιμοποιεί υπολογιστικές τεχνικές, με σκοπό την εκμάθηση, την κατανόηση και την παραγωγή περιεχομένου ανθρώπινης γλώσσας. Οι πρώτες υπολογιστικές προσεγγίσεις στην γλωσσική έρευνα, επικεντρώθηκαν στην αυτοματοποίηση της ανάλυσης της γλωσσικής δομής της γλώσσας και στην ανάπτυξη βασικών τεχνολογιών όπως η μηχανική μετάφραση, η αναγνώριση ομιλίας καθώς και η σύνθεση λόγου. Οι σημερινοί ερευνητές, τελειοποιούν και χρησιμοποιούν τέτοια εργαλεία σε πραγματικές εφαρμογές, δημιουργώντας συστήματα προφορικού διαλόγου, μηχανές μετάφρασης, εξόρυξη δεδομένων που αφορούν θέματα υγείας και οικονομίας σε social media υπηρεσίες, εντοπισμός συναισθημάτων απέναντι σε προϊόντα και υπηρεσίες κλπ.

Τα τελευταία είκοσι χρόνια, η υπολογιστική γλωσσολογία έχει εξελιχθεί τόσο σε έναν συναρπαστικό τομέα επιστημονικής έρευνας, όσο και σε μια πρακτική τεχνολογία που

ενσωματώνεται ολοένα και περισσότερο σε καταναλωτικά προϊόντα όπως για παράδειγμα, το Siri της Apple και το Skype Translator. Οι βασικοί παράγοντες οι οποίοι επέτρεψαν την εξέλιξη αυτή ήταν

- Η τεράστια αύξηση της υπολογιστικής ισχύος
- Η διαθεσιμότητα μεγάλων ποσοτήτων γλωσσικών δεδομένων
- Ανάπτυξη εξαιρετικά επιτυχημένων μεθόδων μηχανικής μάθησης
- Η πλούσια κατανόηση της δομής της ανθρώπινης γλώσσας και της ανάπτυξής της σε κοινωνικά πλαίσια.

Η υπολογιστική γλωσσολογία η οποία είναι γνωστή ως επεξεργασία φυσικής γλώσσας (NLP), είναι το υποπεδίο της επιστήμης των υπολογιστών, που ασχολείται με τη χρήση υπολογιστικών τεχνικών για την εκμάθηση, την κατανόηση και την παραγωγή περιεχομένου ανθρώπινης γλώσσας. Τα υπολογιστικά γλωσσολογικά συστήματα, μπορούν να έχουν πολλαπλούς σκοπούς

- Η υποβοήθηση της επικοινωνίας ανθρώπου με άνθρωπο (μηχανική μετάφραση)
- Η υποβοήθηση επικοινωνίας ανθρώπου με μηχανή (συνομιλητικός πράκτορας - bot)
- Η γενικότερη συνολική ωφέλεια, τόσο των ανθρώπων όσο και των μηχανών από την τεράστια ποσότητα διαθέσιμου περιεχομένου.

Κατά τη διάρκεια των πρώτων δεκαετιών, των τεχνικών, στην υπολογιστική γλωσσολογία, οι επιστήμονες, προσπάθησαν να καταγράψουν τα λεξιλόγια και τους κανόνες των ανθρώπινων γλωσσών για τους υπολογιστές. Αυτό, αποδείχθηκε ένα δύσκολο εγχείρημα λόγω της μεταβλητότητας, της ασάφειας και της εξαρτώμενης από τα συμφραζόμενα ερμηνείας, των ανθρώπινων γλωσσών. Από τη δεκαετία του 80 αλλά ευρύτερα, τη δεκαετία του 90 η NLP μετασχηματίστηκε από τους ερευνητές που άρχισαν να δημιουργούν μοντέλα, πάνω σε μεγάλες ποσότητες εμπειρικών γλωσσικών δεδομένων. Η NLP, με βάση τα στατιστικά ή το "σώμα λέξεων" ήταν μια από τις πρώτες αξιοσημείωτες επιτυχίες της χρήσης μεγάλων δεδομένων, πολύ πριν αναγνωριστεί γενικότερα η δύναμη του Machine Learning και πριν καν εισαχθεί ο όρος, Big Data.

Μία κεντρική διαπίστωση αυτής της στατιστικής προσέγγισης στο NLP ήταν ότι οι απλές μέθοδοι που χρησιμοποιούν λέξεις, ακολουθίες του λόγου (POS) όπως για παράδειγμα αν μία λέξη είναι ουσιαστικό ρήμα ή πρόθεση, μπορούν να πετύχουν αξιοσημείωτα αποτελέσματα όταν εκπαιδεύονται σε μεγάλες ποσότητες δεδομένων. Πολλοί ταξινομητές

κειμένου και συναισθήματος, εξακολουθούν να βασίζονται αποκλειστικά στα διαφορετικά σύνολα λέξεων (bag of words) που περιέχουν τα έγγραφα, χωρίς να λαμβάνουν υπόψη τη δομή ή το νόημα της πρότασης και του λόγου. Η επίτευξη βελτιώσεων σε αυτή την περίπτωση, μπορεί να είναι αρκετά δύσκολη. Τα συστήματα με τις καλύτερες επιδόσεις, πλέον, χρησιμοποιούν εξελιγμένες προσεγγίσεις ML και μια πλούσια κατανόηση της γλωσσικής δομής. Υπάρχουν πλέον διαθέσιμα εργαλεία υψηλής απόδοσης που εντοπίζουν συντακτικές και σημασιολογικές πληροφορίες, καθώς και πληροφορίες σχετικά με το πλαίσιο του λόγου. Ένα τέτοιο παράδειγμα είναι το Stanford CoreNLP το οποίο περιέχει ένα τυπικό pipeline προεπεξεργασίας NLP που περιλαμβάνει

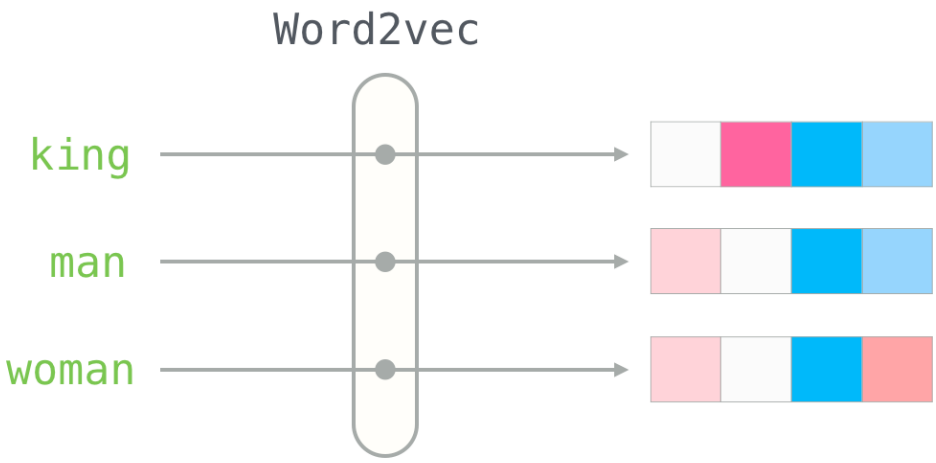
- POS Tagging, με ετικέτες όπως ουσιαστικό, ρήμα και πρόθεση.
- Αναγνώριση ονομαστικών οντοτήτων όπως άνθρωποι, τόποι και οργανισμοί.
- Ανάλυση των προτάσεων στις γραμματικές τους δομές.
- Εντοπισμό coreferences μεταξύ αναφορών ουσιαστικών και φράσεων.

2.3.1 Διανυσματική αναπαράσταση λέξεων

Η διανυσματική αναπαράσταση λέξεων είναι μια τεχνική στην επεξεργασία φυσικής γλώσσας NLP, που αντιστοιχίζει λέξεις ή φράσεις σε μια διανυσματική αναπαράσταση υψηλών διαστάσεων. Με άλλα λόγια, είναι ένας τρόπος μετατροπής λέξεων σε αριθμητικές τιμές που μπορούν να χρησιμοποιηθούν ως είσοδος σε μοντέλα μηχανικής μάθησης. Η διανυσματική αναπαράσταση λέξεων, είναι σημαντική στο NLP επειδή επιτρέπει στους αλγορίθμους μηχανικής μάθησης να επεξεργάζονται τα δεδομένα κειμένου πιο αποτελεσματικά. Αντί οι λέξεις να αντιμετωπίζονται ως μεμονωμένα σύμβολα, η διανυσματική αναπαράσταση λέξεων επιτρέπει στους αλγόριθμους να συλλαμβάνουν το νόημα και τις σχέσεις μεταξύ των λέξεων. Αυτό μπορεί να είναι αρκετά χρήσιμο για ένα ευρύ φάσμα εργασιών NLP όπως της μετάφρασης γλώσσας, της ταξινόμησης κειμένων, της ανάλυσης συναισθημάτων και άλλα.

Το Word2vec για παράδειγμα, είναι μία δημοφιλής τεχνική διανυσματικής αναπαράστασης λέξεων στην επεξεργασία φυσικής γλώσσας και χρησιμοποιεί ένα νευρωνικό δίκτυο για να μάθει διανυσματικές αναπαραστάσεις λέξεων σε ένα σώμα. Ο στόχος του word2vec σύμφωνα με τους Mikolov *et. al. (2013, pp.1-2)* είναι να κατανοήσει το σημασιολογικό νόημα των λέξεων, αναπαριστάνοντας τες ως σταθερού μήκους διανύσματα σε έναν χώρο υψηλών διαστάσεων. Τα διανύσματα αυτά έπειτα, μπορούν να χρησιμοποιηθούν ως είσοδος

σε διάφορες εργασίες NLP, όπως ταξινόμηση κειμένου, ανάλυση συναισθήματος και μετάφραση γλώσσας.

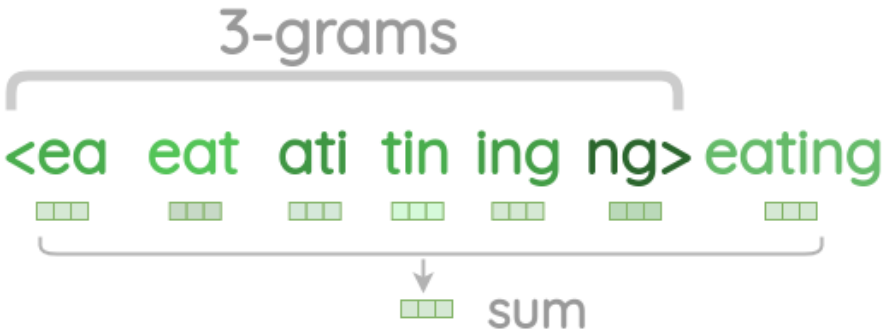


Εικόνα 9: Word2vec

Αποτελείται από δύο μοντέλα. Το μοντέλο Continuous Bag of Words (CBOW) καθώς και το μοντέλο Skip-gram. Το μοντέλο CBOW, παίρνει ένα σύνολο από περιβάλλουσες λέξεις ως είσοδο και προσπαθεί να προβλέψει τη λέξη-στόχο στη μέση, ενώ το μοντέλο Skip-gram παίρνει μία λέξη-στόχο ως είσοδο και προσπαθεί να προβλέψει τις γύρω λέξεις. Και τα δύο μοντέλα εκπαιδεύονται χρησιμοποιώντας έναν αλγόριθμο backpropagation, ο οποίος ενημερώνει τα βάρη του νευρωνικού δικτύου με βάση τη διαφορά μεταξύ των προβλεπόμενων και των πραγματικών εξόδων. Τα διανύσματα λέξεων που προκύπτουν από τα μοντέλα word2vec έχουν αρκετές ιδιότητες. Για παράδειγμα, λέξεις που σημασιολογικά είναι παρόμοιες, τείνουν να έχουν παρόμοιες διανυσματικές αναπαραστάσεις. Αυτό επιτρέπει πράξεις όπως πρόσθεση και αφαίρεση διανυσμάτων, καθιστώντας εύκολη την δημιουργία αναλογιών ή την εύρεση λέξεων με παρόμοια σημασία.

Το FastText είναι μια άλλη δημοφιλής τεχνική ενσωμάτωσης λέξεων, στην επεξεργασία φυσικής γλώσσας και έχει αναπτυχθεί από την ομάδα AI Research (FAIR) του Facebook. Όπως το word2vec, σύμφωνα με τους Bojanovski *et. al. (*2017, pp.135-136*)* το FastText χρησιμοποιεί επίσης ένα νευρωνικό δίκτυο για να μάθει διανυσματικές αναπαραστάσεις λέξεων, σε ένα σώμα. Όμως, το FastText, ενσωματώνει επίσης πληροφορίες υπολέξεων στις αναπαραστάσεις λέξεων, οι οποίες του επιτρέπουν να χειρίζεται λέξεις εκτός λεξιλογίου και να αποτυπώνει το νόημα των τμημάτων της λέξης. Οι πληροφορίες υπολέξεων στο FastText, αναπαρίστανται χρησιμοποιώντας n-gram χαρακτήρες, οι οποίοι

αποτελούν συνεχόμενες ακολουθίες χαρακτήρων μήκους n. Συμπεριλαμβάνοντας n αναπαραστάσεις υπολέξεων στις ενσωματώσεις του, το FastText μπορεί να χειριστεί λέξεις που δεν υπήρχαν στα δεδομένα εκπαίδευσης, καθώς και λέξεις με παρόμοια προθέματα ή επιθήματα.



Εικόνα 10: FastText

Το FastText μπορεί να εκπαιδευτεί σε μεγάλες ποσότητες δεδομένων κειμένου χρησιμοποιώντας μια εποπτευόμενη ή μη εποπτευόμενη προσέγγιση. Στην προσέγγιση χωρίς επίβλεψη, το FastText μαθαίνει ενσωματώσεις λέξεων χρησιμοποιώντας ένα μοντέλο skip-gram παρόμοιο με το word2vec. Στην εποπτευόμενη προσέγγιση, το FastText μπορεί επίσης να χρησιμοποιηθεί για την εκπαίδευση ταξινομητών κειμένου, όπου οι ενσωματώσεις λέξεων χρησιμοποιούνται ως χαρακτηριστικά εισαγωγής.

Ο GloVe είναι ένας χωρίς επίβλεψη αλγόριθμος μάθησης, ο οποίος μαθαίνει τις ενσωματώσεις λέξεων παραγοντοποιώντας τον λογάριθμο του πίνακα συν-εμφάνισης λέξεων του κειμένου. Βασίζεται στην ιδέα ότι μπορούμε να αντλήσουμε σημασιολογικές σχέσεις μεταξύ λέξεων από τα στατιστικά στοιχεία της συνύπαρξής τους σε μεγάλα σώματα κειμένου. Σε αντίθεση με άλλους μεθόδους, το GloVe, έχει σχεδιαστεί για να συλλαμβάνει το νόημα των λέξεων με βάση τις σχέσεις τους με άλλες λέξεις στο σώμα, αντί να εστιάζει απλώς, στο τοπικό πλαίσιο της κάθε λέξης. Το μοντέλο αξιοποιεί αποτελεσματικά τις στατιστικές πληροφορίες εκπαιδεύοντας μόνο τα μη μηδενικά στοιχεία σε έναν πίνακα ταυτόχρονης εμφάνισης λέξης και όχι ολόκληρο τον αραιό πίνακα. Σύμφωνα με τους Pennington *et. al. (2014)* το μοντέλο παράγει έναν διανυσματικό χώρο με ουσιαστική υποδομή, όπως αποδεικνύεται κι από την απόδοση 75% σε μία εργασία αναλογίας λέξεων.

Ορισμένα άλλα μοντέλα είναι τα εξής:

- ELMo

- BERT
- Transformer-XL
- ULMFiT
- ELMo+GloVe
- Contextualized Word Embeddings (CoVe)
- GPT

2.3.2 TF-IDF

Η Term Frequency – Inverse Document Frequency (TF-IDF) είναι μία τεχνική που χρησιμοποιείται στην επεξεργασία φυσικής γλώσσας για να αναπαραστήσει την σημασία ενός όρου, σε ένα σύνολο εγγράφων. Χρησιμοποιείται συχνά για την εξαγωγή σχετικών πληροφοριών και την εκτέλεση ταξινόμησης κειμένου, ομαδοποίησης ή ανάκτησης πληροφοριών. Το TF-IDF χρησιμοποιείται σε ένα σύνολο εγγράφων όταν επιθυμούμε να προσδιορίσουμε τη σημασία των λέξεων ή των όρων σε κάθε έγγραφο, σε σχέση με ολόκληρο το σώμα. Μετρά τη συχνότητα ενός όρου μέσα σε ένα έγγραφο και την κλιμακώνει την αντίστροφη συχνότητα εγγράφου του όρου σε ολόκληρο το σώμα. Αυτό βοηθά να δοθεί μεγαλύτερη βαρύτητα σε όρους που είναι σημαντικοί σε ένα συγκεκριμένο έγγραφο αλλά δεν είναι απαραίτητα κοινοί, σε ολόκληρο το σώμα.

Πιο συγκεκριμένα, σύμφωνα με τους Salton et. al. (1988) κύρια λειτουργία ενός συστήματος στάθμισης όρων αποτελεί η ενίσχυση της αποτελεσματικότητας ανάκτησης. Η αποτελεσματική αυτή ανάκτηση, εξαρτάται από δύο βασικούς παράγοντες. Πρώτο παράγοντα αποτελούν τα στοιχεία που είναι πιθανό να σχετίζονται με τις ανάγκες του χρήστη και πρέπει να ανακτηθούν και δεύτερο παράγοντα αποτελούν τα στοιχεία εκείνα που δεν είναι σχετιζόμενα με τις ανάγκες του χρήστη και οφείλουν να απορριφθούν. Δύο είναι τα μέτρα που χρησιμοποιούνται συνήθως για την αξιολόγηση της ικανότητας ενός συστήματος να πετύχει τα από πάνω και αυτά είναι η ανάκληση και η ακρίβεια. Ανάκληση είναι η αναλογία των σχετικών αντικειμένων που ανακτώνται, μετρούμενη με την αναλογία του αριθμού των σχετικών ανακτημένων αντικειμένων, προς τον συνολικό αριθμό των σχετικών αντικειμένων στη συλλογή. Η ακρίβεια από την άλλη πλευρά είναι η αναλογία των ανακτημένων στοιχείων που είναι σχετικά, μετρούμενη με την αναλογία του αριθμού των σχετικών ανακτημένων στοιχείων προς τον συνολικό αριθμό των ανακτημένων στοιχείων. Γενικότερα, το ιδανικό είναι ένα σύστημα να παράγει τόσο υψηλή ανάκληση, ανακτώντας οτιδήποτε είναι σχετικό, όσο και υψηλή ακρίβεια, απορρίπτοντας όλα τα

στοιχεία που είναι ξένα. Στην πράξη, οι συμβιβασμοί γίνονται συνήθως με την χρήση όρων που είναι αρκετά ευρείς, ώστε να επιτυγχάνεται ένα λογικό επίπεδο ανάκλησης, χωρίς ταυτόχρονα να παράγουν αδικαιολόγητα χαμηλή ακρίβεια.

Οι διαφορετικές απαιτήσεις ανάκλησης και ακρίβεια, ευνοούν τη χρήση σύνθετων συντελεστών στάθμισης όρων που περιέχουν στοιχεία ανάκλησης και βελτίωσης της ακρίβεια. Οι όροι που αναφέρονται συχνά σε μεμονωμένα έγγραφα, φαίνεται να είναι χρήσιμοι ως συσκευές ενίσχυσης της ανάκλησης. Αυτό υποδηλώνει ότι ένας παράγοντας συχνότητας όρου (tf), χρησιμοποιείται ως μέρος του συστήματος στάθμισης όρων, ο οποίος μετρά τη συχνότητας εμφάνισης των όρων στο κείμενο.

$$TF = \frac{TotalWordAppearance}{TotalWords}$$

Οι παράγοντες συχνότητας όρου από μόνοι τους, δεν μπορούν να εξασφαλίσουν αποδεκτή απόδοση ανάκτησης. Συγκεκριμένα, όταν οι όροι υψηλής συχνότητας δεν συγκεντρώνονται σε λίγα συγκεκριμένα έγγραφα, αλλά επικρατούν σε ολόκληρη τη συλλογή, όλα τα έγγραφα τείνουν να ανακτώνται και αυτό επηρεάζει την ακρίβεια αναζήτησης. Έτσι, εισάγεται νέος παράγοντας που εξαρτάται από τη συλλογή και ευνοεί τους όρους που συγκεντρώνονται σε λίγα έγγραφα μιας συλλογής. Αυτό, επιτυγχάνεται με τον παράγοντα αντίστροφης συχνότητας εγγράφου (idf)

$$IDF = log \frac{AllDocumentNumber}{DocumentFrequency}$$

Οι εκτιμήσεις διάκρισης όρων, υποδηλώνουν ότι οι καλύτεροι όροι για την αναγνώριση περιεχομένου του εγγράφου είναι εκείνοι που μπορούν να διακρίνουν ορισμένα μεμονωμένα έγγραφα από το υπόλοιπο της συλλογής. Αυτό σημαίνει ότι οι καλύτεροι όροι, θα πρέπει να έχουν υψηλές συχνότητες όρου αλλά χαμηλές συχνότητες συνολικής συλλογής. Έτσι, μπορεί να ληφθεί ένα εύλογο μέτρο σημασίας όρου, χρησιμοποιώντας το γινόμενο του όρου συχνότητας και της αντίστροφης συχνότητας εγγράφου.

$$TF - IDF = TF\ x\ IDF$$

Η τιμή TF-IDF αυξάνεται αναλογικά με τον αριθμό των φορών που εμφανίζεται μια λέξη στο έγγραφο αλλά αντισταθμίζεται από τη συχνότητα της λέξης στο σώμα, γεγονός που βοηθά στον έλεγχο του γεγονότος ότι ορισμένες λέξεις είναι πιο κοινές από κάποιες άλλες. Η τιμή του TF-IDF κυμαίνεται από το μηδέν μέχρι το ένα, έχοντας δεκαψήφια ακρίβεια. Όσο υψηλότερη είναι η βαθμολογία, τόσο πιο σημαντικός είναι ο όρος. Καθώς λοιπόν ένας όρος γίνεται λιγότερο σχετικός, η βαθμολογία του TF-IDF θα πλησιάζει το 0.

Έστω πως για παράδειγμα, έχουμε τις εξής δύο προτάσεις. «The bike is driven on the roadway» και «The car is driver on the highway». Ο υπολογισμός του TF-IDF θα οδηγήσει στον Πίνακα 2 που ακολουθεί.

Λέξη	**Πρόταση Α**	**Πρόταση Β**	**IDF**	**TF-IDF A**	**TF-IDF B**
the	1/7	1/7	$\log\left(\frac{2}{2}\right) = 0$	0	0
car	1/7	0	$\log\left(\frac{2}{1}\right) = 0.3$	0.043	0
bike	0	1/7	$\log\left(\frac{2}{1}\right) = 0.3$	0	0.043
is	1/7	1/7	$\log\left(\frac{2}{2}\right) = 0$	0	0
driven	1/7	1/7	$\log\left(\frac{2}{2}\right) = 0$	0	0
on	1/7	1/7	$\log\left(\frac{2}{2}\right) = 0$	0	0
the	1/7	1/7	$\log\left(\frac{2}{2}\right) = 0$	0	0
roadway	1/7	0	$\log\left(\frac{2}{1}\right) = 0.3$	0.043	0
highway	0	1/7	$\log\left(\frac{2}{1}\right) = 0.3$	0	0.043

Πίνακας 2: Υπολογισμός TF-IDF

Σε αυτόν, παρατηρούμε και την σημαντικότητα των λέξεων της κάθε πρότασης. Παρατηρούμε ότι οι σημαντικές λέξεις της κάθε πρότασης αντίστοιχα είναι οι bike και roadway για την πρώτη πρόταση και οι λέξεις car και highway για την δεύτερη πρόταση.

2.4 Ταξινόμηση πολλαπλών ετικετών

Σύμφωνα με τον Tsoumakas et. al. (2007) στο παρελθόν, η ταξινόμηση πολλαπλών ετικετών υποκινούταν συνήθως από την ανάγκη κατηγοριοποίησης κειμένων και της ιατρικής διάγνωσης. Κι αυτό επειδή είναι σύνηθες, έγγραφα κειμένου, να ανήκουν σε περισσότερες από μία εννοιολογικές κατηγορίες. Για παράδειγμα μία περίληψη ταινίας, μπορεί να ανήκει και στην κατηγορία Drama αλλά και στην κατηγορία Thriller. Ομοίως, όσον αφορά την ιατρική διάγνωση, ένας ασθενής μπορεί να πάσχει ταυτόχρονα από διαβήτη αλλά και από καρκίνο του προστάτη για παράδειγμα.

Οι μέθοδοι ταξινόμησης, μπορούν να ομαδοποιηθούν σε δύο κύριες κατηγορίες. Τις μεθόδους μετασχηματισμού προβλημάτων καθώς και τις μεθόδους προσαρμογής αλγορίθμων. Οι μέθοδοι μετασχηματισμού προβλήματος, μετατρέπουν το πρόβλημα πολλαπλών ετικετών, σε ένα σύνολο προβλημάτων μίας ετικέτας. Οι μέθοδοι προσαρμογής αλγορίθμων προσαρμόζουν τους υπάρχοντες αλγόριθμους ταξινόμησης μιας ετικέτας ούτως ώστε να χειρίζονται προβλήματα πολλαπλών ετικετών.

2.4.1 Μέθοδοι μετασχηματισμού προβλήματος

Αυτή η μέθοδος, λειτουργεί με την υπόθεση ότι οι αλγόριθμοι ταξινόμησης μιας ετικέτας, είναι πιο κατάλληλοι για την μοντελοποίηση κάθε ετικέτας χωριστά. Το μετασχηματισμένο πρόβλημα, λύνεται χρησιμοποιώντας έναν αλγόριθμο ταξινόμησης μίας ετικέτας και η τελική πρόβλεψη πολλαπλών ετικετών, λαμβάνεται συνδυάζοντας τις εξόδους των ταξινομητών μίας ετικέτας. Μία από τις πιο γνωστές μεθόδους μετασχηματισμού προβλημάτων, είναι η Binary Relevance η οποία σύμφωνα με τους Zhang et. al. (2018) μετατρέπει ένα πρόβλημα πολλαπλών ετικετών με N ετικέτες σε προβλήματα δυαδικής ταξινόμησης N, απλής ετικέτας. Κάθε δυαδικός ταξινομητής, εκπαιδεύεται να προβλέπει την παρουσία ή την απουσία μιας μεμονωμένης ετικέτας και η τελική πρόβλεψη, λαμβάνεται συνδυάζοντας τις εξόδους όλων των δυαδικών ταξινομητών. Μία άλλη κοινή μέθοδος μετασχηματισμού προβλήματος είναι το Label Powerset, το οποίο σύμφωνα με τον

Nareshpalsingh et. al. (2017) μετατρέπει ένα πρόβλημα πολλαπλών ετικετών με N ετικέτες σε ένα πρόβλημα ταξινόμησης πολλαπλών κλάσεων. Σε αυτή τη μέθοδο, κάθε διακριτός συνδυασμός ετικετών, αντιμετωπίζεται ως ξεχωριστή κλάση και ένας ταξινομητής πολλαπλών κλάσεων, εκπαιδεύεται ώστε να προβλέπει μία από αυτές τις κλάσεις. Μία τρίτη μέθοδος μετασχηματισμού προβλήματος είναι οι αλυσίδες ταξινομητή, οι οποίες σύμφωνα με τον Read et. al. (2009) μετατρέπουν ένα πρόβλημα πολλαπλών ετικετών με N ετικέτες, σε προβλήματα δυαδικής ταξινόμησης N απλής ετικέτας. Η ιδέα σε αυτή τη μέθοδο είναι πως κάθε ετικέτα, μπορεί να εξαρτάται από τις προηγούμενες, επομένως, συνδέοντας τους ταξινομητές μεταξύ τους, το μοντέλο μπορεί να λάβει υπόψιν αυτές τις εξαρτήσεις. Για παράδειγμα, εάν ο πρώτος ταξινομητής προβλέπει ότι υπάρχει κάποια συγκεκριμένη ετικέτα, ο επόμενος ταξινομητής μπορεί να χρησιμοποιήσει αυτή την πληροφορία για να βελτιώσει την πρόβλεψή του για την επόμενη ετικέτα. Ένα μειονέκτημα αυτής της μεθόδου είναι πως μπορεί να είναι πολύ απαιτητική σε υπολογιστική ισχύ, διότι κάθε ετικέτα, απαιτεί δικό της ταξινομητή για να εκπαιδευτεί.

2.4.2 Μέθοδοι προσαρμογής αλγορίθμων

Οι μέθοδοι προσαρμογής αλγορίθμων αποτελούν την κατηγορία εκείνη των τεχνικών που χρησιμοποιούνται στη μηχανική εκμάθηση πολλαπλών ετικετών για την προσαρμογή αλγορίθμων ταξινόμησης μιας ετικέτας όσον αφορά τη διαχείριση προβλημάτων πολλαπλών ετικετών. Αυτές οι μέθοδοι, λειτουργούν, τροποποιώντας τον αλγόριθμο ώστε να χειρίζεται πολλαπλές ετικέτες ή συνδυάζοντας πολλαπλές παρουσίες του αλγορίθμου με τρόπο που μπορεί να χειριστεί αποτελεσματικά τα δεδομένα των πολλαπλών ετικετών. Ένα παράδειγμα μιας τέτοιας μεθόδου προσαρμογής αλγορίθμου είναι η Binary Relevance η οποία εκπαιδεύει έναν ξεχωριστό δυαδικό ταξινομητή για κάθε ετικέτα. Ένα άλλο παράδειγμα είναι το Label Powerset το οποίο μετατρέπει το πρόβλημα πολλαπλών ετικετών σε πρόβλημα πολλών κατηγοριών δημιουργώντας έναν μοναδικό συνδυασμό ετικετών για κάθε περίπτωση εκπαίδευσης. Άλλες μέθοδοι προσαρμογής αλγορίθμων περιλαμβάνουν Multi-Label k-Nearest Neighbors (ML-kNN), ο οποίος επεκτείνει τον αλγόριθμο k-Nearest Neighbors για να χειριστεί δεδομένα πολλαπλών ετικετών.

2.4.3 Μετρήσεις αξιολόγησης

Η ταξινόμηση πολλαπλών ετικετών απαιτεί διαφορετικές μετρήσεις από αυτές που χρησιμοποιούνται στην παραδοσιακή ταξινόμηση μιας ετικέτας. Έστω ότι έχουμε D ένα multi-label dataset το οποίο αποτελείται από $|D|$ παραδείγματα πολλαπλών ετικετών $(x_i, Y_i), i = 1..|D|, Y_i \subseteq L$. Έστω το H αποτελεί έναν ταξινομητή πολλαπλών ετικετών και $Z_i = H(x_i)$ να είναι το σύνολο ετικετών που προβλέπονται από το H για παράδειγμα x_i.

Μία συνήθης τεχνική μέτρησης είναι η απώλεια Hamming η οποία μετρά το κλάσμα των ετικετών που έχουν προβλεφθεί λανθασμένα για μία δεδομένη περίπτωση.

$$HammingLoss(H, D) = \frac{1}{|D|}\sum_{i=1}^{|D|}\frac{|Y_i \Delta Z_i|}{|L|}$$

Επιπλέον χρησιμοποιείται η μέτρηση Accuracy η οποία μπορεί να χρησιμοποιηθεί με δύο διαφορετικούς τρόπους. Στον δεύτερο, προστίθεται η παράμετρος $a \geq 0$ η οποία χρησιμοποιείται για την «συγχώρεση» των σφαλμάτων που γίνονται στην πρόβλεψη ετικετών.

$$Accuracy(H, D) = \frac{1}{|D|}\sum_{i=1}^{|D|}\frac{|Y_i \cap Z_i|}{|Y_i \cup Z_i|}$$

$$Accuracy(H, D) = \frac{1}{|D|}\sum_{i=1}^{|D|}\left(\frac{|Y_i \cap Z_i|}{|Y_i \cup Z_i|}\right)^a$$

Επιπλέον, χρησιμοποιούνται οι τεχνικές μέτρησης ακρίβειας (Precision) και ανάκλησης (Recall). Η ακρίβεια, είναι μία μέτρηση η οποία χρησιμοποιείται συνήθως για την αξιολόγηση της απόδοσης των μοντέλων ταξινόμησης πολλών ετικετών. Ορίζεται ως ο λόγος των σωστά προβλεπόμενων θετικών περιπτώσεων, προς τον συνολικό αριθμό των περιπτώσεων που προβλέπονται ως θετικές. Από την άλλη, η ανάκληση, μετρά το ποσοστό των σχετικών ετικετών που έχουν προβλεφθεί σωστά από το μοντέλο. Η ανάκληση κατά

μέσο όρο λαμβάνει υπόψη τη συχνότητα κάθε ετικέτας στο σύνολο δεδομένων και δίνει μεγαλύτερη βαρύτητα στις ετικέτες που εμφανίζονται πιο συχνά.

$$Precision(H, D) = \frac{1}{|D|}\sum_{i=1}^{|D|}\frac{|Y_i \cap Z_i|}{|Z_i|}$$

$$Recall(H, D) = \frac{1}{|D|}\sum_{i=1}^{|D|}\frac{|Y_i \cap Z_i|}{|Y_i|}$$

Στις μετρήσεις αξιολόγησης, υπάρχουν οι όροι micro και macro οι οποίοι αναφέρονται σε διαφορετικές τεχνικές υπολογισμού μέσου όρου, που χρησιμοποιούνται για τον υπολογισμό των από πάνω μετρήσεων όσον αφορά προβλήματα ταξινόμησης πολλαπλών ετικετών. Η micro μέτρηση αξιολόγησης, λαμβάνει υπόψιν τις συνολικές μετρήσεις των αληθινών θετικών, των ψευδώς θετικών και των ψευδών αρνητικών σε όλες τις κατηγορίες και υπολογίζει την μέτρηση απ' όλες αυτές τις συντετρωτικές μετρήσεις. Αντίθετα, η macro τεχνική υπολογίζει τη μέτρηση ανεξάρτητα για κάθε τάξη και στη συνέχεια, παίρνει τον μέσο όρο σε όλες τις κατηγορίες και είναι κατάλληλο για την αξιολόγηση της απόδοσης του ταξινομητή όσον αφορά μεμονωμένες κατηγορίες δίχως να λαμβάνει υπόψη την ανισορροπία κλάσης σε αντίθεση με την micro τεχνική, η οποία είναι κατάλληλη όταν επιθυμούμε μέτρηση συνολικής απόδοσης. Σύμφωνα με τον Mendsaikhan et. al. (2020) αυτοί είναι οι μαθηματικοί τύποι για τις micro και macro μετρήσεις.

$$\text{P}recision = \frac{\sum_{j=1}^{Q} TP_j}{\sum_{j=1}^{Q} TP_j + \sum_{j=1}^{Q} FP_j}$$

$$\text{Recall} = \frac{\sum_{j=1}^{Q} TP_j}{\sum_{j=1}^{Q} TP_j + \sum_{j=1}^{Q} FN_j}$$

Μετά την εύρεση των micro Precision και Recall, μπορούμε να υπολογίσουμε και το F1 Score, το οποίο είναι μία αρμονική μέση τιμή των δύο αυτών τιμών.

$$\mathrm{F1} = \frac{2\ x\ microPrecision\ x\ microRecall}{microPrecision\ + microRecall}$$

Αντίστοιχα, όσον αφορά τα macro Precision και Recall έχουμε τους εξής τύπους.

$$Precision = \frac{1}{Q}\sum_{j=1}^{Q}\frac{TP_j}{TP_j + FP_j}$$

$$Recall = \frac{1}{Q}\sum_{j=1}^{Q}\frac{TP_j}{TP_j + FN_j}$$

Μετά την εύρεση των macro Precision και Recall, μπορούμε να υπολογίσουμε και το F1 Score, το οποίο είναι μία αρμονική μέση τιμή των δύο αυτών τιμών.

$$F1 = \frac{1}{Q}\sum_{j=1}^{Q}\frac{2\ x\ P_j\ x\ R_j}{P_j + R_j}$$

2.5 Επαύξηση δεδομένων

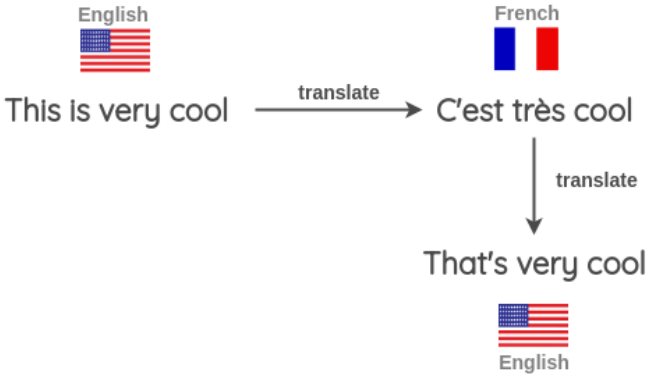
Η επαύξηση δεδομένων είναι μία τεχνική η οποία χρησιμοποιείται για την αύξηση της ποικιλομορφίας και του μεγέθους ενός συνόλου δεδομένων. Ο τρόπος με τον οποίο γίνεται είναι με την δημιουργία νέων δειγμάτων χρησιμοποιώντας τα υπάρχοντα. Η τεχνική αυτή χρησιμοποιείται ευρέως στις τεχνικές deep learning για την βελτίωση της γενίκευσης και της ανθεκτικότητας των μοντέλων μηχανικής μάθησης. Όσον αφορά της εργασίες ταξινόμησης κειμένου, η επαύξηση δεδομένων περιλαμβάνει την δημιουργία καινούργιων δειγμάτων κειμένου, τα οποία είναι παρόμοια με τα αρχικά, έχοντας απλά μικρές παραλλαγές στο περιεχόμενο, τη δομή και το στυλ τους. Αυτό επιτυγχάνεται μέσω

διαφόρων τεχνικών όπως η παράφραση, η αναδρομική μετάφραση, η αντικατάσταση λέξεων, η σύνθεση κειμένου κλπ.

Η επαύξηση δεδομένων αποδεικνύεται ιδιαίτερα αποτελεσματική για τη βελτίωση στην απόδοση των μοντέλων ταξινόμησης κειμένου, ιδίως στα σενάρια εκείνα στα οποία, τα διαθέσιμα δεδομένα με ετικέτα είναι περιορισμένα ή ανισοροπημένα. Με τη δημιουργία αυτών των νέων δειγμάτων κειμένου, το μοντέλο, μπορεί να μάθει διαφορετικά και πιο αντιπροσωπευτικά τα μοτίβα στα δεδομένα, μειώνοντας έτσι την υπερπροσαρμογή και βελτιώνοντας ταυτόχρονα την ακρίβεια των προβλέψεων. Επιπλέον, το μοντέλο με τα καινούργια αυτά δείγματα, μπορεί να προσαρμοστεί πιο εύκολα σε δείγματα τα οποία έχουν διαφορετικά χαρακτηριστικά από τα δεδομένα εκπαίδευσης.

2.5.1 Αναδρομική μετάφραση

Η αναδρομική μετάφραση, όπως περιγράφουν και οι Sennrich et all (2016) είναι μια τεχνική στην επεξεργασία της φυσικής γλώσσας, που περιλαμβάνει την μετάφραση ενός κειμένου από μία γλώσσα σε μία άλλη και στην συνέχεια, μετάφραση ξανά στην αρχική. Ο πρωταρχικός στόχος της, είναι η δημιουργία πρόσθετων δεδομένων εκπαίδευσης, για μοντέλα μηχανικής μάθησης, ιδιαίτερα στον τομέα της μηχανικής μετάφρασης. Έχει αποδειχθεί ιδιαιτέρως αποτελεσματική στην αύξηση της ποικιλομορφίας των δεδομένων εκπαίδευσης. Η διαδικασία της μετάφρασης μπορεί να πραγματοποιηθεί πολλές φορές, δημιουργώντας συνεχώς ελαφρώς διαφορετικές προτάσεις από την αρχική επιτρέποντας έτσι στα μοντέλα μηχανικής εκμάθησης, να μαθαίνουν από ποικίλο σύνολο δεδομένων, οδηγώντας σε καλύτερη απόδοση.

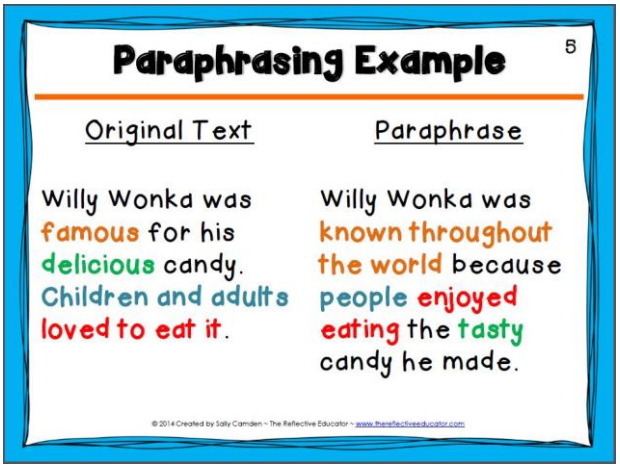


Εικόνα 11: Αναδρομική μετάφραση

Υπάρχουν διάφορα εργαλεία για την αναδρομική μετάφραση. Το OpenNMT για παράδειγμα, επιτρέπει στους χρήστες να εκπαιδεύουν τα δικά τους μοντέλα αυτόματης μετάφρασης και να τα χρησιμοποιούν για τεχνικές αναδρομικής μετάφρασης. Επίσης, το Moses είναι μία άλλη εργαλειοθήκη ανοιχτού κώδικα για στατιστική, αυτόματη μετάφραση και προσαρμόζεται εύκολα σε συγκεκριμένες απαιτητικές εργασίες. Τέλος, το Hugging Face Transformers αποτελεί κι αυτό μια δημοφιλή βιβλιοθήκη ανοιχτού κώδικα για τον NLP και περιλαμβάνει διάφορα προ-εκπαιδευμένα μοντέλα για αυτόματη μετάφραση. Μπορεί να χρησιμοποιηθεί για αναδρομική μετάφραση, δημιουργώντας μεταφράσεις μιας πρότασης κειμένου σε διαφορετικές γλώσσας και έπειτα, μεταφράζοντας πίσω στην επιθυμητή.

2.5.2 Παράφραση

Η παράφραση, σύμφωνα και με τους Zhou et al. (2021) αποτελεί μία τεχνική που χρησιμοποιείται συνήθως για την αύξηση του μεγέθους ενός συνόλου δεδομένων, μέσω της δημιουργίας επιπρόσθετων παραδειγμάτων. Στον πλαίσιο της επεξεργασίας φυσικής γλώσσας (NLP), η παράφραση περιλαμβάνει την αναδιατύπωση μιας πρότασης ή ενός κομματιού κειμένου, διατηρώντας όμως παράλληλα το αρχικό της νόημα. Και με αυτή τη μέθοδο, αυξάνεται η ποικιλομορφία του συνόλου δεδομένων οδηγώντας πιθανά, σε καλύτερη απόδοση ενός μοντέλου.



Εικόνα 12: Παράδειγμα παράφρασης

Η τεχνική αυτή, είναι ιδιαίτερα χρήσιμη όταν το μέγεθος του συνόλου είναι περιορισμένο ή όταν υπάρχει η λεγόμενη ανισορροπία. Παράδειγμα παράφρασης αποτελεί το ανακάτεμα

των προτάσεων, που περιλαμβάνει το ανακάτεμα της σειρά των λέξεων ή των φράσεων σε μια πρόταση, διατηρώντας παράλληλα το αρχικό νόημα. Επίσης, έχουμε την διαίρεση και συγχώνευση προτάσεων, που περιλαμβάνει τον διαχωρισμό μίας πρότασης ή τη συγχώνευση δύο ή περισσότερων. Μία ενδιαφέρον τεχνική παράφρασης αποτελεί ή «άρνηση και επιβεβαίωση» που περιλαμβάνει την αλλαγή της πολικότητας μίας πρότασης, με την προσθήκη ή την αφαίρεση λέξεων όπως όχι ή ποτέ. Ένα παράδειγμα θα μπορούσε να είναι η πρόταση «Μου αρέσει η πίτσα» η οποία θα μπορούσε να παραφραστεί ως «Δεν μισώ την πίτσα»

2.5.3 Τυχαία διαγραφή λέξεων

Η τυχαία διαγραφή όπως περιγράφεται κι από τους Wei et al (2019) είναι μία τεχνική αύξησης κειμένου που περιλαμβάνει την τυχαία αφαίρεση λέξεων από μία πρόταση ή από ένα έγγραφο. Η τεχνική αυτή, έχει αποδειχθεί αποτελεσματική στην δημιουργία νέων δεδομένων εκπαίδευσης για μοντέλα μηχανικής μάθησης, όσον αφορά τις διάφορες εργασίες επεξεργασίας φυσικής γλώσσας, όπως ταξινόμηση κειμένου, ανάλυση συναισθήματος κλπ. Η τυχαία διαγραφή, μπορεί να βοηθήσει στη δημιουργία διαφορετικών δεδομένων εκπαίδευσης και να βελτιώσει την ανθεκτικότητα των μοντέλων, στις παραλλαγές της γλώσσας.

Εικόνα 13: Τυχαία διαγραφή λέξης

3. Σύνολο δεδομένων

3.1 Κατάσταση ευπαθειών 2018/19

Η έκθεση η οποία δημοσίευσε ο Οργανισμός της Ευρωπαϊκής Ένωσης για την Κυβερνοασφάλεια (ENISA) παρέχει μία εκτενή ανάλυση της κατάστασης των ευπαθειών λογισμικού και της εκμετάλλευσής τους στην Ευρωπαϊκή Ένωση, κατά τη διάρκεια του 2018 και του 2019. Επισημαίνεται από τους Katos *et. al*. (2019, pp.5-8) πως ο αριθμός των ευπαθειών που αναφέρθηκαν στην ΕΕ, συνέχισε να αυξάνεται και οι ευπάθειες εφαρμογών ιστού, είναι οι πιο συχνά αναφερόμενες, αντιπροσωπεύοντας περισσότερο του 50% όλων των αναφερόμενων ευπαθειών. Επιπλέον, υπογραμμίζεται ο αυξανόμενος αριθμός ευπαθειών που εντοπίζεται σε συσκευές Internet of Things (IoT) με σημαντικό αριθμό αυτών, να αφορά συσκευές απλών καταναλωτών.

Επισημαίνεται επίσης πως, η εκμετάλλευση των ευπαθειών παραμένει μία σημαντική απειλή, με τους επιτιθέμενους να χρησιμοποιούν συχνά γνωστές ευπάθειες για να αποκτήσουν μη εξουσιοδοτημένη πρόσβαση σε συστήματα και δεδομένα. Σημειώνεται πως ο μέσος χρόνος επιδιόρθωσης των ευπαθειών παραμένει σχετικά μεγάλος με κάποιες από τις ευπάθειες, να παραμένουν δίχως επιδιόρθωση ακόμα και για χρόνια μετά την εύρεσή τους. Η σημασία της προληπτικής προσέγγισης των οργανισμών όσον αφορά τη διαχείριση των ευπαθειών και την εφαρμογή αποτελεσματικών διαδικασιών διαχείριση των ευπαθειών για την μείωση του κινδύνου εκμετάλλευσης, είναι πολύ μεγάλη. Συνιστώνται διάφορες βέλτιστες πρακτικές για τη διαχείριση των ευπαθειών, συμπεριλαμβανομένου της χρήσης αυτοματοποιημένων εργαλείων και μία τακτική επιδιόρθωσης ήδη γνωστών ευπαθειών. Υπογραμμίζεται επίσης η ανάγκη βελτίωσηw της συνεργασίας και της ανταλλαγής σημαντικών πληροφοριών μεταξύ των διαφόρων ενδιαφερόμενων μερών, δηλαδή των ερευνητών ασφαλείας, των εταιρειών λογισμικού καθώς και των τελικών χρηστών.

3.2 Ανάλυση δεδομένων

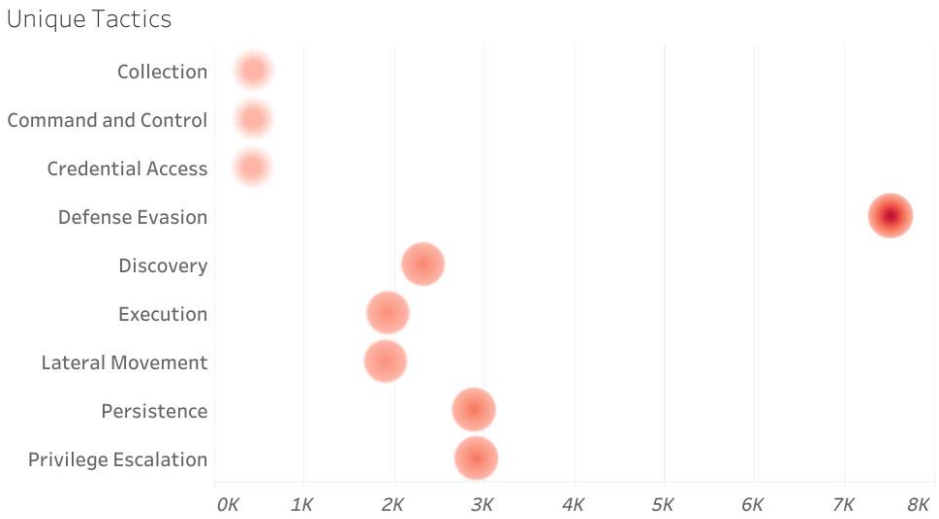
Η απόκτηση των δεδομένων έγινε από διάφορες πηγές όπως, βάσεις δεδομένων ευπαθειών και ταξινομήσεις στο NVD Database, στο ATT&CK, Shodan, Exploit-db κλπ. Έγινε επίσης από άρθρα, τα οποία αφορούν συγκεκριμένες ευπάθειες προσφέροντας περισσότερες λεπτομέρειες για αυτές, από αδόμητα δεδομένα προερχόμενα από αναφορές των εταιρειών λογισμικού και διάφορες πηγές γενικών ειδήσεων και διάφορα άλλα. Μέρος της ανάλυσης

των δεδομένων που συλλέχθηκαν, αποτελούσε η χαρτογράφηση της κάθε ευπάθειας στις αντίστοιχες τεχνικές MITRE ATT&CK χρησιμοποιώντας τις κοινές απαριθμήσεις και ταξινομήσεις μοτίβων επίθεσης (CAPEC) που βρέθηκαν τόσο στο NVD όσο και στο ATT&CK. Για τη συγκέντρωση των απαιτούμενων δεδομένων, όλες οι πηγές μελετήθηκαν και αξιολογήθηκαν, προκειμένου να διασφαλιστεί ότι τα δεδομένα αυτά είναι κατάλληλα, καθώς και υψηλής ποιότητας. Η βασική απαίτηση για την επιλογή της κάθε πηγής δεδομένων ήταν πως, τα δεδομένα αυτά θα πρέπει να παρέχονται δωρεάν και κυρίως σε μία μορφή δομημένου συνόλου δεδομένων.

Συλλέχθηκαν συνολικά 27.471 πληροφορίες ευπάθειας, οι οποίες είχαν δημοσιευτεί από την 1η Ιανουαρίου 2018 έως και τις 30 Σεπτεμβρίου 2019. Οι πληροφορίες, ανάμεσα σε άλλα περιλαμβάνουν χαρακτηριστικά όπως τα εξής: το cvss3 score και severity, οι τεχνικές και οι τακτικές στο MITRE Attack, ένα κείμενο περιγραφής της κάθε ευπάθειας, την ημερομηνία δημοσίευσης, το όνομα του προϊόντος καθώς και της εταιρίας υλικολογισμικού.

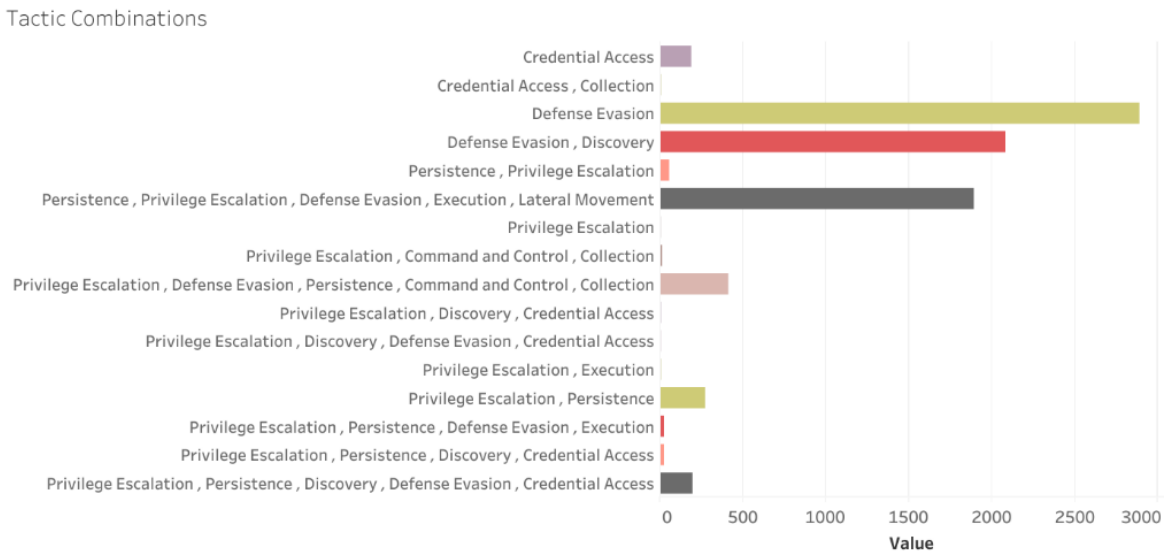
3.2.1 Τακτικές MITRE

Χρησιμοποιώντας όλα τα διαθέσιμα δεδομένα και μελετώντας τα, όπως φαίνεται και στην Εικόνα 17, παρατηρούμε πως η πιο συνηθισμένη τακτική, δηλαδή ο λόγος που ο επιτιθέμενος πραγματοποιεί την ενέργεια, είναι το Defense Evasion και ακολουθούν το Privilege Escalation και το Persistence.



Εικόνα 14: Μοναδικές τακτικές

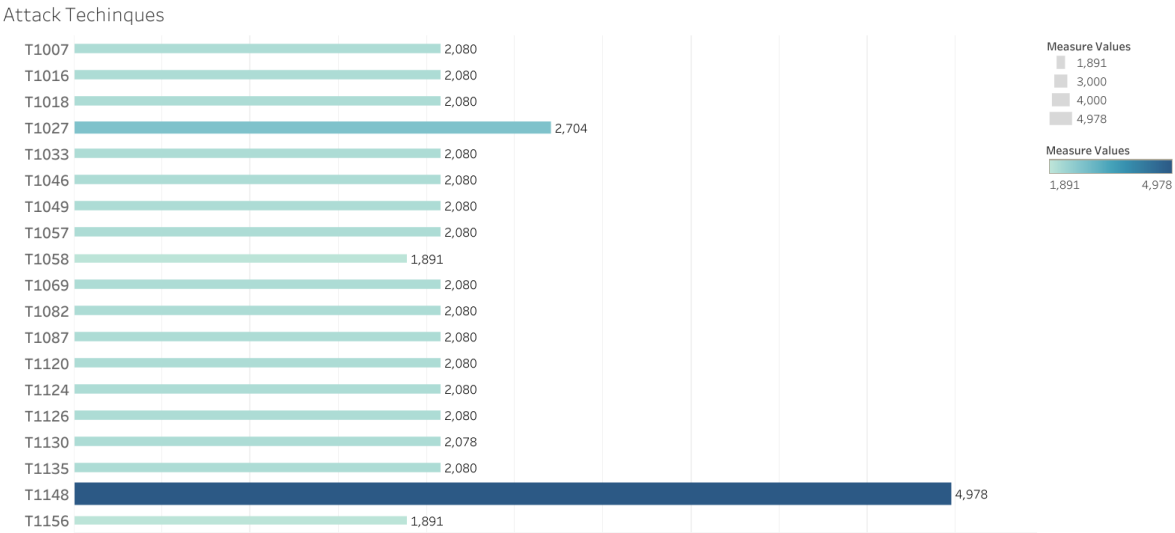
Παρατηρείται επίσης πως η τακτική του Defense Evasion κυριαρχεί στα διαθέσιμα δεδομένα μας και εκτός από μόνη της εμφανίζεται στους πιο συχνά επαναλαμβανόμενους συνδυασμούς όπως στο Discovery - Defense Evasion και στον συνδυασμό Persistence-Privilege Escalation – Defense Evasion – Execution – Lateral Movement. Στην Εικόνα 18 φαίνονται όλοι οι διαθέσιμοι συνδυασμοί στα δεδομένα μας.



Εικόνα 15: Συνδυασμοί Τακτικών

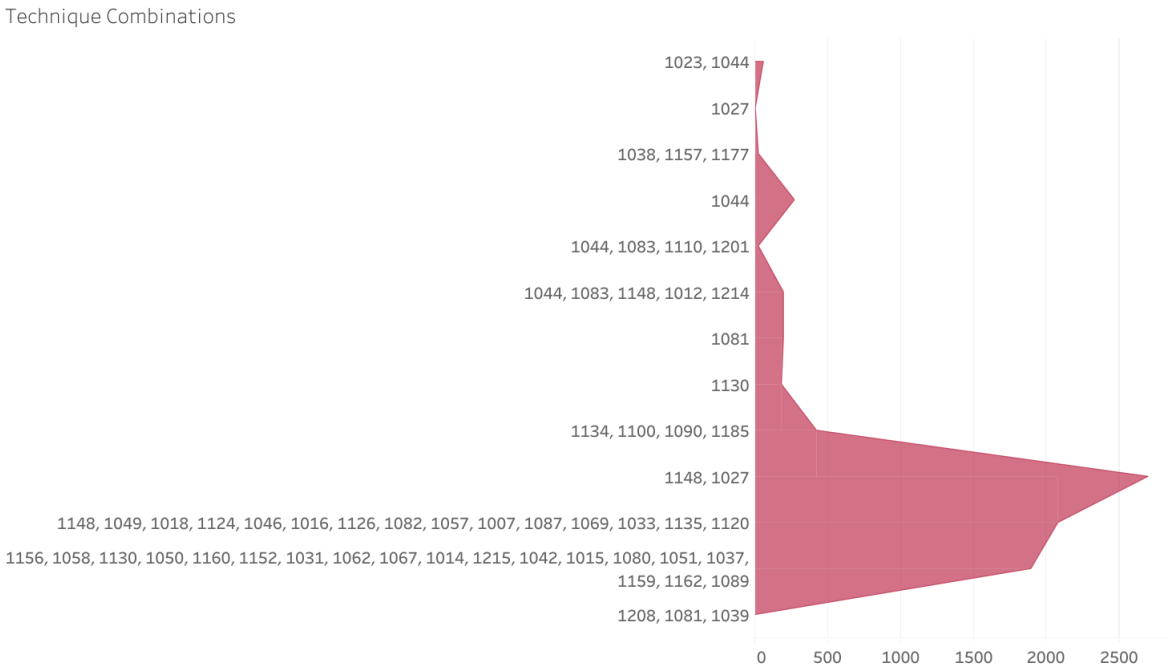
3.2.2 Τεχνικές MITRE

Όσον αφορά τις τεχνικές MITRE, η T1148 επαναλαμβάνεται 4978 φορές, η T1027 επαναλαμβάνεται 2704 φορές. Στην Εικόνα 19 παρουσιάζονται οι πιο συχνά επαναλαμβανόμενες τεχνικές στα δεδομένα μας.



Εικόνα 16: Τεχνικές Mitre

Στους συνδυασμούς των τεχνικών που εμφανίζονται στην κάθε ευπάθεια, ξεχωρίζει ο συνδυασμό των T1148 – T1027 ακολουθούμενος από δύο διαφορετικούς συνδυασμούς τεχνικών οι οποίοι περιλαμβάνουν δώδεκα και δεκαεννέα μοναδικές τεχνικές αντίστοιχα. Όλοι οι συνδυασμοί εμφανίζονται στην Εικόνα 20.



Εικόνα 17: Συνδυασμοί Τεχνικών

3.2.3 Ανάλυση κειμένου

Χρησιμοποιώντας ξανά, το σύνολο των διαθέσιμων δεδομένων και συγκεκριμένα όσον αφορά το description, που περιλαμβάνει κείμενο με την πλήρη περιγραφή της ευπάθειας βλέπουμε τα εξής στατιστικά χαρακτηριστικά όπως παρουσιάζονται στον Πίνακα 3.

Μέσο μήκος χαρακτήρων	304.58
Median μήκος χαρακτήρων	242.00
Ελάχιστο μήκος χαρακτήρων	23
Μέγιστο μήκος χαρακτήρων	3761

Πίνακας 3: Συνοπτικά στατιστικά στοιχεία

Χρησιμοποιώντας μία γραφική αναπαράσταση των λέξεων που εμφανίζονται πιο συχνά στο description, δημιουργούμε το λεγόμενο «σύννεφο λέξεων» για τον γρήγορο εντοπισμό των σημαντικότερων εξ' αυτών. Για ουσιαστικότερο αποτέλεσμα, αφαιρέθηκαν τα λεγόμενα «stop words», λέξεις δηλαδή όπως "the", "we" κλπ. Στην Εικόνα 21 φαίνεται το σύννεφο αυτό.

Εικόνα 18: Σύννεφο Λέξεων

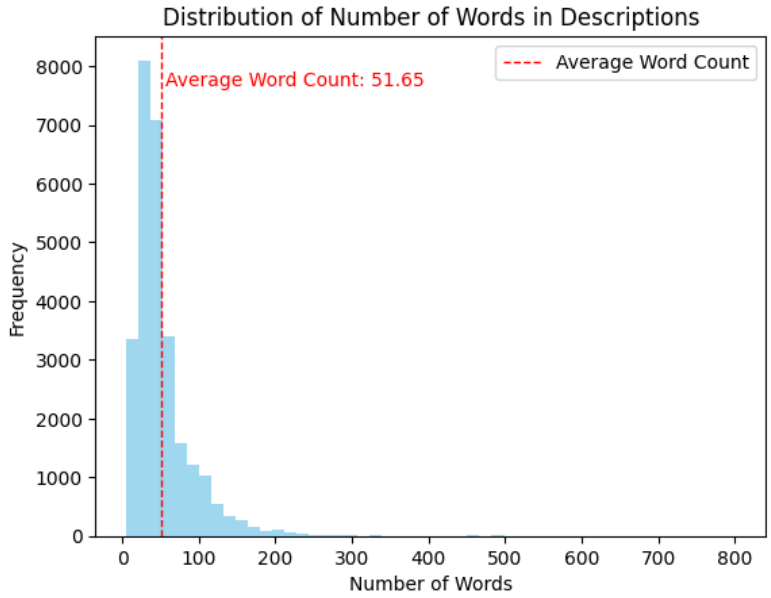
Με μία γρήγορη ματιά λοιπόν, παρατηρούμε ότι στο σύνολο των description, εμφανίζονται συνεχώς λέξεις όπως "discovered", "attacker", "remote", "vulnerability" κλπ. Πιο συγκεκριμένα, στον Πίνακα 4 καταγράφονται οι έξι πιο συχνά επαναλαμβανόμενες λέξεις των descriptions.

vulnerability	14740
attacker	10211
allows	8329
via	7779
could	7387
remote	6429

Πίνακας 4: Συνολική Εμφάνιση Λέξεων

Όσον αφορά τον συνολικό αριθμό των λέξεων παρατηρούμε ότι η πλειοψηφία των δεδομένων μας, περιλαμβάνει descriptions τα οποία φτάνουν περίπου τις 150 λέξεις με μέσο

αριθμό λέξεων να είναι λίγο πάνω από τις 51 λέξεις ανά description όπως φαίνεται και στην Εικόνα 22.



Εικόνα 19: Αριθμός λέξεων

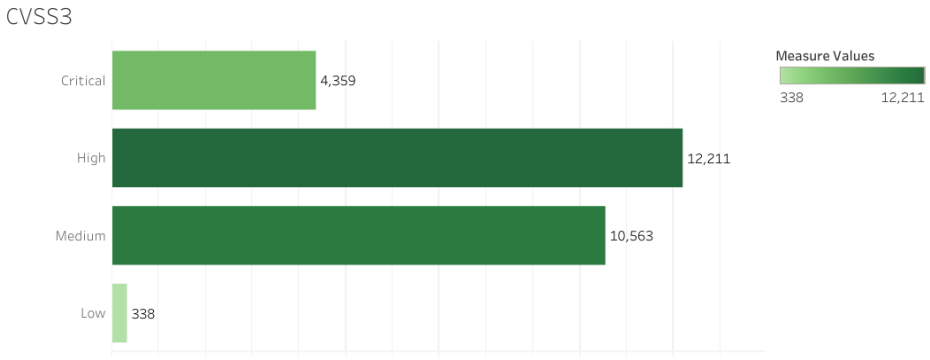
3.2.3 Ανάλυση CVSS3

Το CVSS είναι ένα πλαίσιο το οποίο παρέχει έναν τρόπο τυποποίησης της μέτρησης της σοβαρότητας των τρωτών σημείων ασφαλείας. Η βαθμολογία της είναι μία αριθμητική τιμή που κυμαίνεται από το 0 έως το 10, με το 10 να αποτελεί την πιο σοβαρή ευπάθεια. Υπολογίζεται με βάση διάφορες μετρήσεις, όπως ο αντίκτυπος στην εμπιστευτικότητα, την ακεραιότητα και τη διαθεσιμότητα, την πολυπλοκότητα που απαιτείται για την εκμετάλλευση καθώς και το επίπεδο της απαιτούμενης αλληλεπίδρασης με τον χρήστη. Η κατηγοριοποίηση με βάση την βαθμολογία, καθορίζεται ως εξής όπως φαίνεται και στον Πίνακα 5.

Critical	9.0-10.0
High	7.0-8.9
Medium	4.0-6.9
Low	0.1-3.9

Πίνακας 5: Κατηγοριοποίηση βαθμολογιών

Αναλύοντας όλα τα διαθέσιμα δεδομένα μας και όπως φαίνεται στην Εικόνα 23, παρατηρούμε ότι η πλειοψηφία των ευπαθειών βρίσκεται στις κατηγορίες High – Medium με τρίτη στη σειρά να είναι η Critical και με πολύ μικρότερο ποσοστό, να εμφανίζονται ευπάθειες κατηγορίας Low.



Εικόνα 20: Κατηγοριοποίηση CVSS3

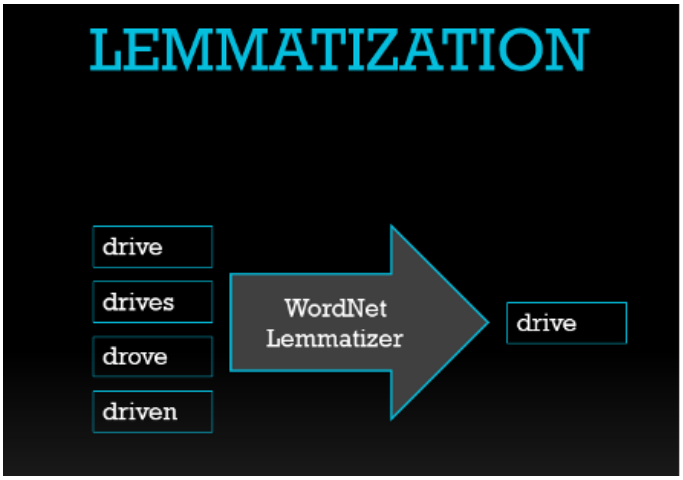
4. Αποτελέσματα

Για το πρόβλημα ταξινόμησης ετικετών, χρησιμοποιήθηκε το χαρακτηριστικό “description” και το χαρακτηριστικό “technique_id”. Στο χαρακτηριστικό “description” υπάρχει η αναλυτική περιγραφή της ευπάθειας με κείμενο και στο “technique_id” υπάρχει η αριθμητική περιγραφή της τεχνικής, σε αντιστοιχία του ATT&CK framework. Σε κάθε περιγραφή, αντιστοιχούν από 1 έως και 19 τακτικές, συνδυαστικά. Από τις συνολικά 27.471 εγγραφές, μόνο οι 8077 από αυτές, έχουν διαθέσιμα δεδομένα όσον αφορά το χαρακτηριστικό “technique_id” το οποίο και θέλουμε να προβλέψουμε. Οι διαφορετικές τεχνικές οι οποίες εμφανίζονται στα δεδομένα μας είναι 58, εκ των οποίων οι έντεκα από αυτές από 1 έως και 59 φορές στο σύνολο των 8077, δηλαδή, λιγότερες από το 1% και επιλέχθηκε να αγνοηθούν, για να μην επηρεάσουν τα αποτελέσματα.

Όπως σε όλα τα προβλήματα μηχανικής μάθησης, πραγματοποιείται η “προετοιμασία” των δεδομένων. Η προετοιμασία συμβαίνει στο χαρακτηριστικό “description” και αφορά τα εξής.

- Διαγραφή των λεγομένων “stop words”.
- Διαγραφή των αριθμών
- Διαγραφή των ειδικών χαρακτήρων
- Διαγραφή των λέξεων που αποτελούνται από ένα γράμμα
- Λημματοποίηση

Η λημματοποίηση, διαδραματίζει κρίσιμο ρόλο στην επεξεργασία φυσικής γλώσσας (NLP) επιτρέποντας πιο ακριβή και ουσιαστική ανάλυση του κειμένου. Στις εργασίες NLP, οι λέξεις, πρέπει να αντιμετωπίζονται με συνέπεια για να «συλλαμβάνεται» η εγγενείς σημασία τους. Η λημματοποίηση βοηθά στην επίτευξη αυτού του στόχου, μειώνοντας τις λέξεις στη βασική τους μορφή ή λήμμα, που επιτρέπει την καλύτερη σύγκριση και κατανόηση του κειμένου. Μετατρέποντας τις λέξεις στα λήμματά τους, μπορούν να επιλυθούν παραλλαγές λόγω χρόνου ή πληθυντικού αριθμού για παράδειγμα, διασφαλίζοντας έτσι πως διαφορετικές μορφές της ίδιας λέξεις, αναγνωρίζονται ως πανομοιότυπες.



Εικόνα 21: Λημματοποίηση

Η λημματοποίηση επιτυγχάνεται με την χρήση του WordNet Lemmatizer. Το WordNet, είναι ένας λεξικολογικός πόρος που ομαδοποιεί τις λέξεις σε σύνολα συνωνύμων που ονομάζονται «synsets» και παρέχει σημασιολογικές σχέσεις μεταξύ τους. Το WordNet Lemmatizer χρησιμοποιεί αυτόν ακριβώς τον πόρο, για να προσδιορίσει το λήμμα ή τη βασική μορφή μιας λέξης. Ακολουθεί ένα παράδειγμα τυχαίου επιλεγμένου «description» πριν και μετά την υλοποίηση όλων των από πάνω βημάτων στην επεξεργασία του.

Αρχικό κείμενο

lib/backup/cli/utility.rb in the backup-agoddard gem 3.0.28 and (2) lib/backup/cli/utility.rb in the backup_checksum gem 3.0.23 for Ruby place credentials on the openssl command line, which allows local users to obtain sensitive information by listing the process.

Επεξεργασμένο κείμενο

lib backup cli utility rb backup agoddard gem lib backup cli utility rb backup checksum gem ruby place credential openssl command line allows local user obtain sensitive information listing process

Ακολούθως, η υλοποίηση του TF-IDF πραγματοποιήθηκε χρησιμοποιώντας τα εξής χαρακτηριστικά

- max_df = 0.92
- max_features = 60000
- ngram_range = (1,3)

Η παράμετρος max_df στο TfIdfVectorizer, επιτρέπει τον καθορισμό της μέγιστης συχνότητας εγγράφου για έναν όρο. Η συχνότητα αυτή αναφέρεται στον αριθμό των εγγράφων σε μία συλλογή, που περιέχουν έναν συγκεκριμένο όρο. Ορίζοντας ένα max_df, μπορούν να εξαιρεθούν όροι που είναι πολύ συνηθισμένοι και ενδέχεται να μην περιέχουν εκείνες τις πληροφορίες που θα προκαλέσουν την επιθυμητή διάκριση.

Η παράμετρος max_features στο TfIdfVectorizer, καθορίζει τον μέγιστο αριθμό χαρακτηριστικών όρων, που θα διατηρηθούν. Έτσι, περιορίζεται το μέγεθος του λεξιλογίου σε εκείνους οι οποίοι θεωρούνται «κορυφαίοι» και σημαντικοί και αυτοί θα συμπεριληφθούν στον προκύπτον πίνακα.

Η παράμετρος ngram_range, επιτρέπει τον καθορισμό της ελάχιστης και της μέγιστης τιμής για το n, ορίζοντας το εύρος των μεγεθών ngram που θα συμπεριληφθούν. Το ελάχιστο όριο αντιπροσωπεύει τον ελάχιστο αριθμό διαδοχικών λέξεων για να σχηματιστεί ένα n-gram και το μέγιστο, αντιπροσωπεύει τον μέγιστο αριθμό διαδοχικών λέξεων. Για παράδειγμα, ngram_range (1,3) σημαίνει ότι επιλέγονται μονές λέξεις, ζευγάρια γειτονικών λέξεων καθώς και ακολουθίες τριών γειτονικών λέξεων.

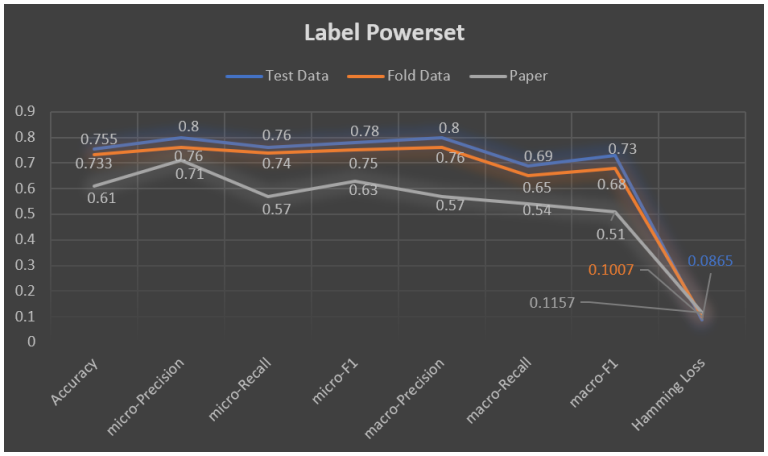
Τα δεδομένα διαχωρίζονται σε Training και Testing, με test_size να αποτελεί το 30% των συνολικά 8.077 δεδομένων μας. Ο διαχωρισμός, χρησιμοποιεί την μέθοδο stratify. Στην μηχανική μάθηση, το «stratify», αναφέρεται σε μια τεχνική που χρησιμοποιείται κατά τη διαδικασία διαχωρισμού των δεδομένων. Όταν τα δεδομένα είναι «στρωματοποιημένα», σημαίνει ότι η κατανομή κλάσης της μεταβλητής στόχου, διατηρείται σε κάθε διαχωρισμό. Όταν για παράδειγμα υπάρχει ένα σύνολο δεδομένων όπου μία κατηγορία είναι σημαντικά πιο διαδεδομένη από τις υπόλοιπες, αυτό σημαίνει ότι δεν υπάρχει ομοιόμορφη κατανομή. Σε αυτές τις περιπτώσεις, είναι σημαντικό να διασφαλιστεί ότι η κατανομή της κλάσης, διατηρείται σε διαφορετικά υποσύνολα δεδομένων για να αποφευχθούν μεροληπτικές αξιολογήσεις της απόδοσης του μοντέλου.

Στον Πίνακα 6, παρουσιάζονται τα αποτελέσματα των μοντέλων. Τα αποτελέσματα αυτά, είναι δίχως την πραγματοποίηση τεχνικών επαύξησης δεδομένων και αφορούν το Testing dataset. Σε παρένθεση, παρουσιάζονται τα διαθέσιμα αποτελέσματα των Mendsaikhan et. al. (2020). Παρατηρείται πως σε όλες τις περιπτώσεις, πετυχαίνουμε καλύτερα αποτελέσματα.

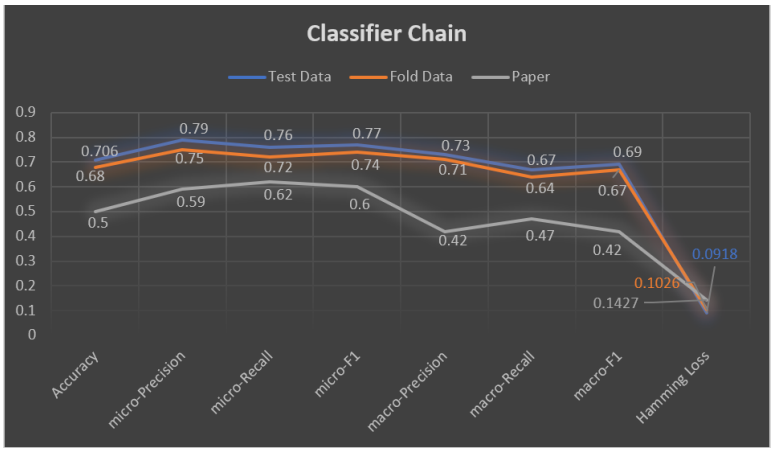
		Micro Average			Macro Average			
Algorithm	Accuracy	Precision	Recall	F1 Score	Precision	Recall	F1 Score	Hamming loss
Label Powerset	0.75 (0.61)	0.80 (0.71)	0.76 (0.57)	0.78 (0.63)	0.8 (0.57)	0.69 (0.54)	0.73 (0.51)	0.0865 (0.1157)
Classifier Chain	0.706 (0.50)	0.79 (0.59)	0.76 (0.62)	0.77 (0.60)	0.73 (0.42)	0.67 (0.47)	0.69 (0.42)	0.0918 (0.1427)
Binary Relevance	0.651 (0.374)	0.76 (0.57)	0.77 (0.62)	0.76 (0.61)	0.69 (0.46)	0.73 (0.61)	0.69 (0.49)	0.0971 (0.1471)
LP (neural)	0.761 (0.743)	0.84 (0.75)	0.76 (0.73)	0.80 (0.74)	0.77 (0.68)	0.70 (0.62)	0.73 (0.63)	0.078 (0.091)
kNN	0.640 (0.613)	0.77 (0.73)	0.67 (0.62)	0.72 (0.67)	0.75 (0.65)	0.60 (0.50)	0.65 (0.50)	0.1076 (0.1079)

Πίνακας 6: Αρχικά αποτελέσματα

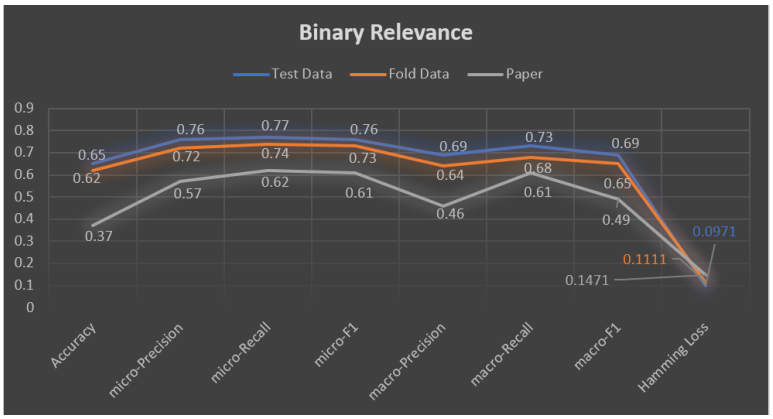
Για πιο ολοκληρωμένη εικόνα, πραγματοποιήθηκε και τεχνική cross_validate στα δεδομένα μας και στα επόμενα σχήματα, παρατίθεται ο μέσος όρος των μετρήσεων αυτών εν συγκρίση με τα αποτελέσματα του Test Data καθώς και με τα αποτελέσματα των Mendsaikhan et. al. (2020).



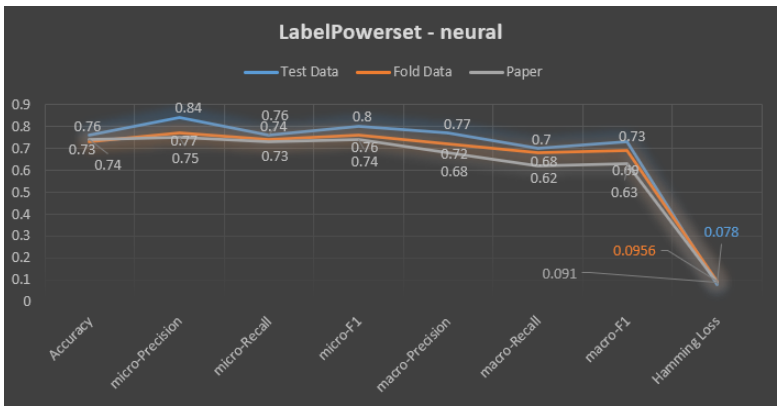
Σχήμα 13: Αποτελέσματα Label Powerset



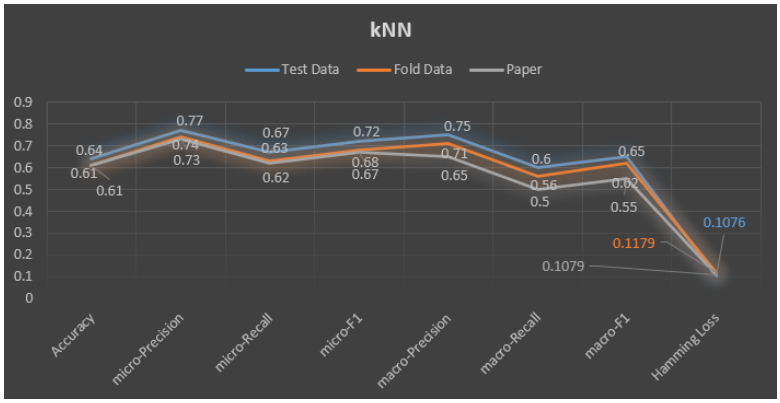
Σχήμα 14: Αποτελέσματα Classifier Chain



Σχήμα 15: Αποτελέσματα Binary Relevance



Σχήμα 16: Αποτελέσματα Neural - LP



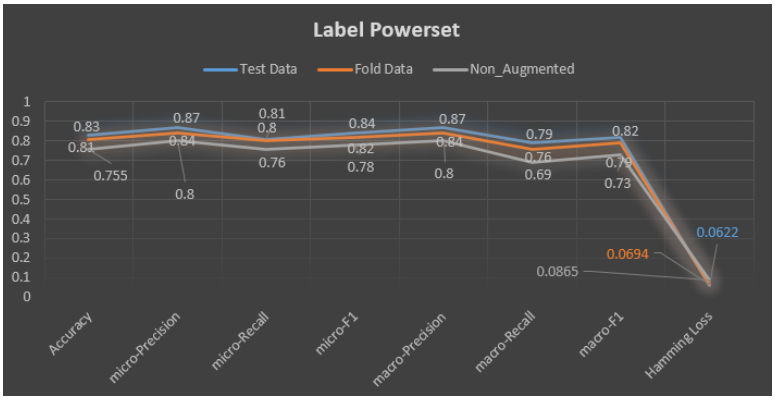
Σχήμα 17: Αποτελέσματα kNN

Στη συνέχεια παρουσιάζονται τα αποτελέσματα των μοντέλων, έπειτα από την εφαρμογή της τεχνικής της αναδρομικής μετάφρασης στα training δεδομένα μας. Η σύγκριση που ακολουθεί, αφορά μόνο με τα αποτελέσματα των δεδομένων πρωτού την πραγματοποίηση τεχνικής επαύθξησης δεδομένων.

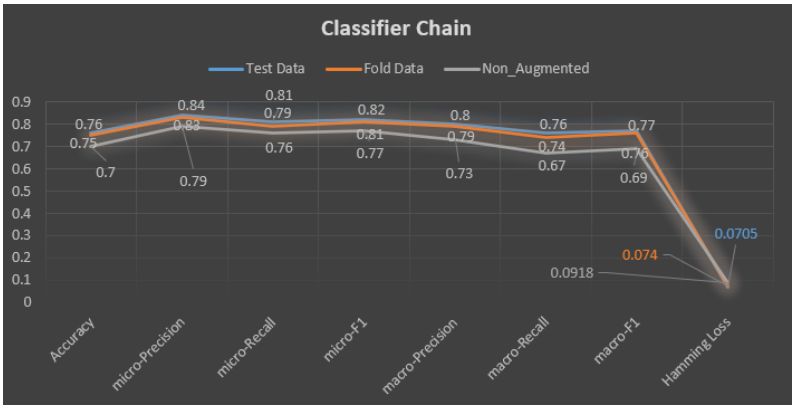
		Micro Average			Macro Average			
Algorithm	Accuracy score	Precision	Recall	F1 Score	Precision	Recall	F1 Score	Hamming loss
LabelPowerset	0.83 (0.755)	0.87 (0.80)	0.81 (0.76)	0.84 (0.78)	0.87 (0.80)	0.79 (0.69)	0.82 (0.73)	0.0622 (0.0865)
ClassifierChain	0.76 (0.706)	0.84 (0.79)	0.81 (0.76)	0.82 (0.77)	0.80 (0.73)	0.76 (0.67)	0.77 (0.69)	0.0705 (0.0918)
BinaryRelevance	0.75 (0.65)	0.85 (0.76)	0.83 (0.77)	0.84 (0.76)	0.80 (0.69)	0.81 (0.73)	0.80 (0.69)	0.063 (0.0971)
LP - neural	0.78 (0.761)	0.8 (0.84)	0.8 (0.76)	0.8 (0.80)	0.77 (0.77)	0.77 (0.70)	0.77 (0.73)	0.0793 (0.078)
kNN	0.69 (0.640)	0.77 (0.77)	0.71 (0.67)	0.74 (0.72)	0.74 (0.75)	0.66 (0.60)	0.70 (0.65)	0.1016 (0.1076)

Πίνακας 7: Augmented αποτελέσματα

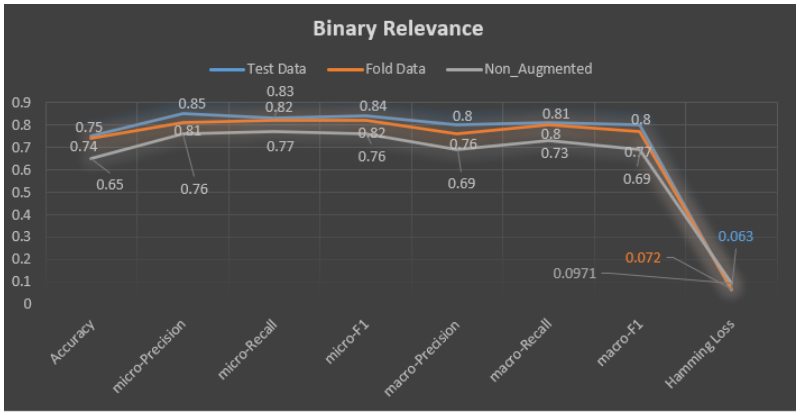
Όπως και στην προηγούμενη περίπτωση, πραγματοποιήθηκε και τεχνική cross_validate στα δεδομένα για μία πιο ολοκληρωμένη εικόνα των αποτελεσμάτων. Οι μετρήσεις αυτές, επιβεβαιώνουν με την σειρά τους τα θετικότερα αποτελέσματα στα μοντέλα μας έπειτα από την χρησιμοποίηση τεχνικών επαύξησης δεδομένων.



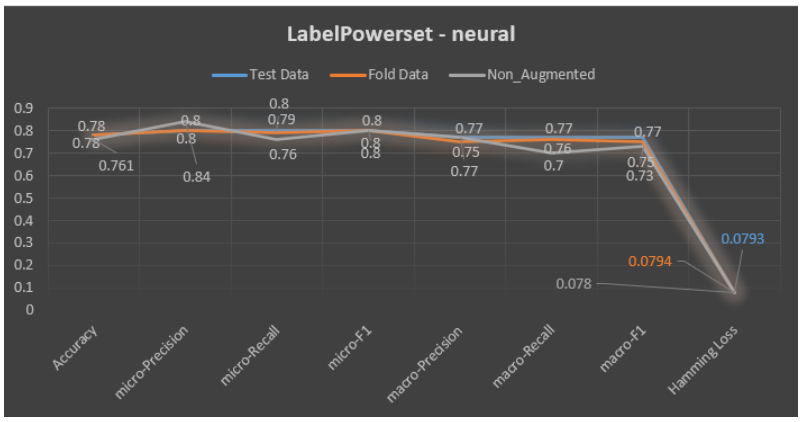
Σχήμα 18: Augmented αποτελέσματα Label Powerset



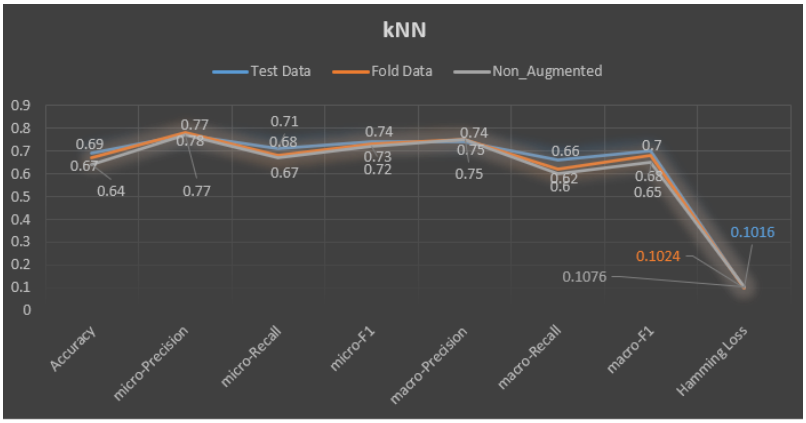
Σχήμα 19: Augmented αποτελέσματα Classifier Chain



Σχήμα 20: Augmented αποτελέσματα Binary Relevance



Σχήμα 21: : Augmented αποτελέσματα Label Powerset - neural



Σχήμα 22: Augmented αποτελέσματα kNN

Βιβλιογραφία

Gollman, D. (2011) Computer security, Hamburg: John Wiley & Son s, Inc.

Blake, S., Applebaum, A., Miller, P., Nickels, K., Pennington, A., Thomas, C., (2020) "MITRE ATT&CK: Design and Philosophy" *MITRE PRODUCT*, 2020, pp.1-4

Katos, V., Rostami, S., Bellonias, P., Davies, N., Kleszcz, A., Faily, S., Spyros, A., Papanikolaou, A., Ilioudis. C., Rantos, K. (2019) 'State of the vulnerabilities 2018/2019', pp.5-8

Ευπάθεια ασφαλείας (2023) *Vulnerabilities, Exploits, and Threats - Vulnerabilities, Exploits, and Threats*.) Ανακτήθηκε από https://www.rapid7.com/fundamentals/vulnerabilities-exploits-threats/ (Πρόσβαση την 4η Μαΐου 2023)

Canadian Centre for Cyber Security (2019) *An Introducation to the Cyber Threat Environment.* Ανακτήθηκε από https://icclr.org/wp-content/uploads/2019/05/Intro-to-cyber-threat-environment-e.pdf?x37853

Cyber Security Statistics - *Cyber security vulnerability statistics and facts of 2022.* Ανακτήθηκε από https://www.comparitech.com/blog/information-security/cybersecurity-vulnerability-statistics/ (Πρόσβαση την 4η Μαΐου 2023)

Mikolov, T., Chen, K., Corrado, G., Dean, J., (2013) 'Efficient Estimation of Word Representations in Vector Space', 7 September, pp.1-2

Bojanowski, P., Grave, E., Joulin, A., Mikolov, T., (2017) 'EnrichingWord Vectors with Subword Information', 19 June, pp.135-136.

Pennington, J., Socher, R., Manning, C., (2014) 'GloVe: Global Vectors forWord Representation' *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1532–1543, Doha, Qatar. Association for Computational Linguistics

ENISA (2023) *About ENISA - The European Union Agency for Cybersecurity.* Διαθέσιμο από: https://www.enisa.europa.eu/about-enisa (Πρόσβαση την 1η Μαΐου 2023)

OWASP (2023) *OWASP Top Ten* OWASP Top 10 Project. Διαθέσιμο από https://owasp.org/www-project-top-ten/ (Πρόσβαση την 5η Μαΐου 2023)

NIST (2023) Vulnerability Metrics National Vulnerability Database. Διαθέσιμο από https://nvd.nist.gov/vuln-metrics/cvss (Πρόσβαση την 5η Μαΐου 2023)

Jordan, M., Mitchell, T., (2015) 'Machine learning: Trends, perspectives, and prospects' Science, Volume 349, pp. 255-257

Naqa, I., Li, R., Murphy, M., (2015) Machine Learning in Radiation Oncology, 1st ed., Springer International Publishing AG Switzerland

Murphy, D. (2012) Machine Learning A Probabilistic Perspective The MIT Press Cambridge, Massachusetts London, England

Kaelbling, L., Littman, M., Moore, A., (1996) Reinforcement Learning A Survey Journal of Artificial Intelligence Research. pp. 237-238

Samuel, A., (1957) Some Studies in Machine Learning Using the Game of Checkers, IBM Journal, Vol. 3, No.3

Rosenblatt, F., (1958). The perceptron: A probabilistic model for information storage and organization in the brain. Psychological Review, 65(6), 386–408.

Werbos, P., (1974). Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Science. Thesis (Ph. D.). Appl. Math. Harvard University.

Quinlan, J.R. (1986*)* Induction of Decision Trees. Machine Learning, 1, 81-106.

Cortes, C., Vapnik, V. (1995). Support-vector networks. Mach Learn 20, 273–297

Shapire, E., (1999). A brief introduction to boosting. In Proceedings of the 16th international joint conference on Artificial intelligence - Volume 2 (IJCAI'99). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1401–1406.

Breiman, L., 2001. *Random forests*. Machine learning, 45, pp.5-32

Hinton, G., Learning multiple layers of representation. *Trends Cogn Sci*. 2007 Oct;11(10):428-34.

Sennrich, R., Haddow, B. and Birch, A., 2015. Improving neural machine translation models with monolingual data. arXiv preprint arXiv:1511.06709.

Zhou, J., Bhat, S., (2021). Paraphrase Generation: A Survey of the State of the Art. *In Proceedings of the 2021 Conference on Empirical Methods in Natural Language*

Processing, pages 5075–5086, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

Salton, G., Buckley, C., (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, Volume 24, Issue 5, Pages 513-523

Wei, J., Kai, Z., (2019). EDA: Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks. *In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 6382–6388, Hong Kong, China. Association for Computational Linguistics.

Hirschberg, J., & Manning, C. D. (2015). *Advances in natural language processing*. Science (New York, N.Y.), 349(6245), 261–266.

Shevchenko, N. (2018). Threat Modeling: 12 Available Methods. Διαθέσιμο από https://insights.sei.cmu.edu/blog/threat-modeling-12-available-methods/. (Πρόσβαση την 6η Μαΐου 2023)

Mahesh, B., 2020. Machine learning algorithms-a review. *International Journal of Science and Research (IJSR). [Internet]*, 9, pp.381-386.

Rish, I., 2001, August. An empirical study of the naive Bayes classifier. *In IJCAI 2001 workshop on empirical methods in artificial intelligence* (Vol. 3, No. 22, pp. 41-46).

Guo, G., Wang, H., Bell, D., Bi, Y. and Greer, K., 2003. KNN model-based approach in classification. *In On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE: OTM Confederated International Conferences, CoopIS, DOA, and ODBASE 2003*, Catania, Sicily, Italy, November 3-7, 2003. Proceedings (pp. 986-996). Springer Berlin Heidelberg.

Maind, S.B. and Wankar, P., 2014. Research paper on basic of artificial neural network. *International Journal on Recent and Innovation Trends in Computing and Communication*, 2(1), pp.96-100.

Maulud, D. and Abdulazeez, A.M., 2020. A review on linear regression comprehensive in machine learning. *Journal of Applied Science and Technology Trends*, 1(4), pp.140-147.

Hosmer Jr, D.W., Lemeshow, S. and Sturdivant, R.X., 2013. Applied logistic regression (Vol. 398). John Wiley & Sons.

Kotsiantis, S.B., 2013**.** Decision trees: a recent overview. *Artificial Intelligence Review*, 39, pp.261-283.

Tsoumakas, G. and Katakis, I., 2007. Multi-label classification: An overview. *International Journal of Data Warehousing and Mining (IJDWM)*, 3(3), pp.1-13.

Zhang, M.L., Li, Y.K., Liu, X.Y. and Geng, X., 2018. Binary relevance for multi-label learning: an overview. *Frontiers of Computer Science*, 12, pp.191-202.

Nareshpalsingh, J.M. and Modi, H.N., 2017. Multi-label classification methods: a comparative study. *International Research Journal of Engineering and Technology (IRJET)*, *4*(12), pp.263-270.

Read, J., Pfahringer, B., Holmes, G. and Frank, E., 2011. Classifier chains for multi-label classification. Machine learning, 85, pp.333-359.

Mendsaikhan, O., Hasegawa, H., Yamaguchi, Y. and Shimada, H., 2020. Automatic mapping of vulnerability information to adversary techniques. In The Fourteenth International Conference on Emerging Security Information, Systems and Technologies SECUREWARE2020.