ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

Μαθηματική Επισκόπηση του Μοντέλου της Λογιστικής Παλινδρόμησης με Εφαρμογή στην Ανίχνευση των Fake News

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΓΚΑΝΕΤΣΟΥ ΣΠΥΡΟΥ

Επιβλέπων : Στεφανέας Πέτρος

Αναπληρωτής Καθηγητής Ε.Μ.Π.

Αθήνα, Σεπτέμβριος 2021

Ευχαριστίες

Ευχαριστώ τον καθηγητή μου κ. Πέτρο Στεφανέα για την ενθάρρυνση και τη βοήθεια που μου προσέφερε κατά τη διάρκεια εκπόνησης της διπλωματικής μου εργασίας, καθώς και τους γονείς μου για την υλική και ψυχολογική στήριξη για την εισαγωγή και την περάτωση των σπουδών μου στη σχολή Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Σκοπός της παρούσας διατριβής είναι η εμβάθυνση στο πρόβλημα της διάδοσης Fake News στο σύγχρονο κόσμο, αναλύοντας τις κοινωνικοπολιτικές προεκτάσεις του και τις τεχνολογικές εξελίξεις που εντείνουν το πρόβλημα αυτό, καθώς και η μαθηματική επισκόπηση και εφαρμογή του μοντέλου της λογιστικής παλινδρόμησης για την ανίχνευση Fake News. Συγκεκριμένα , αναδεικνύονται τα χαρακτηριστικά και τα κίνητρα διάδοσης των Fake News, όπως και οι σημαντικές συνέπειες τους στη δημοκρατία, στα ατομικά δικαιώματα και στην ίδια την ανθρώπινη ζωή. Παράλληλα, αντικείμενο μελέτης αποτελούν οι τεχνολογικές εξελίξεις στον τομέα της τεχνητής νοημοσύνης που έχουν δημιουργήσει νέους τρόπους διάδοσης Fake News αλλά και οι δυνατότητες αξιοποίησης αυτών των τεχνολογιών για την κατάλληλη αντιμετώπιση των Fake News. Τέλος, αναλύεται μαθηματικά το λογιστικό μοντέλο παλινδρόμησης και η εφαρμογή του στην πρόβλεψη του δυαδικού αποτελέσματος ταξινόμησης κειμένων ως αληθινά ή Fake News. Η εφαρμογή και η εκπαίδευση το μοντέλου υλοποιείται σε Γλώσσα Προγραμματισμού Python, αξιοποιώντας δύο σύνολα δεδομένων αποτελούμενα από 23.489 άρθρα Fake News το ένα και 21.418 αληθινά άρθρα το άλλο. Συμπερασματικά, κρίνεται ότι το μοντέλο είναι αρκετά αποδοτικό μιας και κατέγραψε ποσοστό ακρίβειας στις προβλέψεις του 98.81%.

Abstract

This dissertation aims to shed light on the issue of spreading Fake News in the modern world by analyzing its socio-political implications and technological developments that intensify the problem, as well as on the mathematical overview and application of the logistic regression model towards detecting Fake News. More specifically, efforts were made to showcase the features and motives of Fake News along with their significant impact on democracy, individual rights as well as on human life itself. At the same time, emphasis was placed on technological advances in the field of Artificial Intelligence; particularly on how these advances have generated new ways of spreading Fake News and on what opportunities to make use of new technologies exist in order to address Fake News. Finally, the logistic regression model was analyzed mathematically, focusing on how it can be applied to predict the binary text sorting result as real or Fake News. The model was applied and trained with the use of Python Programming Language, utilizing two data sets consisting of 23.489 Fake News articles and 21.418 real ones respectively. In conclusion, it is estimated that the model is fairly efficient, since it recorded a precision of 98.81%.

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΙΝΑΚΑΣ ΕΙΚΟΝΩΝ

ΠΙΝΑΚΑΣ ΓΡΑΦΗΜΑΤΩΝ

ΚΕΦΑΛΑΙΟ 1
ΕΙΣΑΓΩΓΗ

1.1. Τεχνολογία και Fake News

Καθώς ο κόσμος γύρω μας αλλάζει με πολύ γρήγορους ρυθμούς και αναμφίβολα βαδίζουμε προς μια τέταρτη βιομηχανική επανάσταση, παρατηρούμε τις καθημερινές μας συνήθειες να μεταβάλλονται και αυτές. Πλέον μέσω του διαδικτύου μπορούμε να πραγματοποιούμε τις καθημερινές αγορές μας, να πληρώνουμε λογαριασμούς και ανάλογες οικονομικές υποχρεώσεις, να εκτελούμε μεταφορές χρημάτων σε άλλους λογαριασμούς και φυσικά να ενημερωνόμαστε. Μπορεί οι παραπάνω διαδικασίες να μοιάζουν αρκετά οικείες, όμως θα πρέπει να αναλογιστεί κανείς, ότι με πολλές από αυτές 10-15 χρόνια πριν, όχι απλά δεν υπήρχε εξοικείωση στον πληθυσμό, αλλά ήταν άγνωστες και συχνά αντιμετωπίζονταν ως ιδέες παρμένες από ταινίες επιστημονικής φαντασίας.

Σίγουρα η είσοδος των νέων τεχνολογιών στη ζωή μας έχει διευκολύνει σε πολλές πτυχές της την καθημερινότητα μας. Συγκεκριμένα στο κομμάτι της ενημέρωσης, παραδοσιακά μέσα όπως οι εφημερίδες δεν ανήκουν πλέον στις κύριες προτιμήσεις μας, γιατί μέσω του διαδικτύου μπορούμε να έχουμε ελεύθερη και δωρεάν (τις περισσότερες φορές) πρόσβαση στην είδηση και μάλιστα σε ζωντανή ροή. Παράλληλα, μας δίνεται η δυνατότητα να επιλέξουμε εμείς τη στιγμή της ημέρας που θα ενημερωθούμε με βάση τις ανάγκες μας και το πρόγραμμα μας κάτι που αποτελεί συγκριτικό πλεονέκτημα ακόμα και με άλλα μέσα όπως η τηλεόραση που έχει ενημερωτικές εκπομπές και δελτία ειδήσεων προκαθορισμένες ώρες της ημέρας.

Ωστόσο, η πρόσβαση στην εύκολη και γρήγορη είδηση σύντομα δημιούργησε άλλα προβλήματα. Συγκεκριμένα, παρατηρούμε τα τελευταία χρόνια την παραπληροφόρηση στην ενημέρωση να γιγαντώνεται όλο και περισσότερο. Αξιοσημείωτο αποτελεί το γεγονός ότι η παραπληροφόρηση δεν αποτελεί νέο φαινόμενο στα μέσα ενημέρωσης. Αρκετές εφημερίδες, για παράδειγμα, είχαν υιοθετήσει αυτήν την τακτική ανάδειξης εντυπωσιακών ψευδεπίγραφων ειδήσεων (Fake News) με στόχο φυσικά τις περισσότερες πωλήσεις. Άλλοτε πάλι, παρατηρούνταν χρωματισμός των ειδήσεων λόγω σύνδεσης των εκάστοτε ΜΜΕ με οικονομικά ή πολιτικά συμφέροντα.

Το τοπίο σήμερα στην εποχή του διαδικτύου παραμένει σχεδόν ίδιο, αυτό που αλλάζει είναι τα μέσα. Η πραγματικότητα είναι ότι το διαδίκτυο είναι ένα πρόσφορο έδαφος για παραπληροφόρηση μιας και ο οποιοσδήποτε μπορεί είτε ανώνυμα είτε και επώνυμα, αλλά χωρίς το βάρος της ιδιότητας του, να διαδώσει εύκολα και γρήγορα μια άποψη γεμάτη ανακρίβειες ή και ψευδεπίγραφες ειδήσεις. Συγχρόνως, οι τεχνολογικές εξελίξεις στον κλάδο της Τεχνητής Νοημοσύνης καθιστούν τη διάδοση των Fake News ακόμα πιο εύκολη και πιο επικίνδυνη, μιας και δημιουργούνται συνεχώς νέοι μηχανισμοί παραγωγής και διάδοσης τους έτσι ώστε να είναι ολοένα και πιο δύσκολο για το μέσο χρήστη να τα αναγνωρίσει.

1.2. Ιστορική Χρήση

Οι διαστάσεις που έχει λάβει σήμερα το ζήτημα της διάδοσης Fake News στο διαδίκτυο και τα μέσα κοινωνικής δικτύωσης το καθιστούν αναμφισβήτητα ένα κοινωνικό-

πολιτικό ζήτημα. Στις Αμερικάνικες προεδρικές εκλογές το 2016 και στο δημοψήφισμα για το «Brexit» στο Ηνωμένο Βασίλειο, είχαμε αποδεδειγμένα οργανωμένες προσπάθειες μαζικής χειραγώγησης των πολιτών. Μάλιστα, κατά την προεκλογική περίοδο στην Αμερική χρησιμοποιήθηκαν ακόμα και προσωπικά δεδομένα από το Facebook (σκάνδαλο Cambridge Analytica). Σκοπός της κλοπής δεδομένων ήταν η διάδοση των Fake News να γίνεται με έναν πιο στοχευμένο τρόπο και συγκεκριμένα με ειδικούς τύπους μηνυμάτων που καθιστούν κάποιον πολίτη πιο ευάλωτο στα Fake News, με βάση το ψυχολογικό του προφίλ, όπως αυτό κατασκευάστηκε από τα προσωπικά δεδομένα.

Παράλληλα, τα Fake News αξιοποιούνται και διαδίδονται από επιστημονικούς κύκλους που κινούνται στα όρια της ψευδοεπιστήμης και στόχος τους είναι η αυτοπροβολή τους ή η εξυπηρέτηση συγκεκριμένης πολιτικής ή οικονομικής ατζέντας. Με αυτόν τον μηχανισμό έχουν καταφέρει να κερδίσουν έδαφος ακραίες απόψεις όπως ο αντιεμβολιασμός ή ακόμα και ότι η γη είναι επίπεδη (flat earthers).

1.3. Σκοπός και Αντιμετώπιση

Αντιλαμβανόμαστε, λοιπόν, πως πλέον η αναχαίτιση των Fake News αποτελεί όχι απλά ζήτημα δημοκρατίας αλλά και βασική προϋπόθεση για τη διατήρηση της κοινωνικής συνοχής, ακόμα και για την προστασία της ίδιας της ανθρώπινης ζωής. Αυτή λοιπόν η νέα πραγματικότητα, δημιουργεί και νέες προκλήσεις στον επιστημονικό κόσμο να βρεθούν αποδοτικότεροι και γρηγορότεροι τρόποι να ελέγχεται η αξιοπιστία των ειδήσεων. Ήδη μέσα στα τελευταία χρόνια πολλοί αλγόριθμοι έχουν φέρει σημαντικά αποτελέσματα προς αυτή την κατεύθυνση, με άλλους να είναι πιο αποδοτικούς και άλλους λιγότερο. Αυτό βεβαίως εξαρτάται από τον τρόπο που θα επιλέξει ο κάθε ερευνητής να προσεγγίσει το πρόβλημα.

Η παρούσα διπλωματική εργασία έχει σκοπό την ανάδειξη του προβλήματος που δημιουργείται από το φαινόμενο των Fake News καθώς επίσης και να αναφέρει τρόπους και μεθοδολογίες κατά τις οποίες μπορούν αυτά να εντοπιστούν στον τομέα της πληροφόρησης μέσω γραπτών κειμένων. Το κεφάλαιο 2 πραγματεύεται με την ανάλυση του όρου, τις μορφές που εμφανίζεται, καθώς επίσης και τα κίνητρα και τις προθέσεις δημιουργίας τους. Στο κεφάλαιο 3 αναφέρονται με μεγαλύτερη ανάλυση και εμβύθιση οι τεχνολογίες αυτές στις διαφορετικές μορφές τους, θέτοντας το υπόβαθρο για την δυνατότητα αναγνώρισης και αντιμετώπισης αυτών. Το κεφάλαιο 4 αναφέρεται στις μαθηματικές μελέτες παλινδρόμησης που υλοποιούν τις παραπάνω δράσεις και την αντιμετώπιση αυτών, θέτοντας την απαραίτητη μαθηματική μελέτη για κατανόηση των αλγορίθμων. Το κεφάλαιο 5 παρουσιάζει τις ερευνητικές αποφάσεις και αναλύσεις που πραγματοποιήθηκαν για την εκπόνηση της διπλωματικής με τα σχετικά ευρηματικά στοιχεία. Τέλος, το κεφάλαιο 6 παρουσιάζει τα συμπεράσματα και γίνεται η απαραίτητη συζήτηση για τις μελλοντικές εργασίες που απαιτούνται για την βελτίωση των σχετικών αλγορίθμων διαχωρισμού των κειμένων σε αυθεντικά ή ψεύτικα.

ΚΕΦΑΛΑΙΟ 2
ΤΟ ΦΑΙΝΟΜΕΝΟ ΤΩΝ FAKE NEWS

2.1. Misinformation, Disinformation και Fake News

Οι μελέτες σχετικά με τα λεγόμενα fake news έχουν αποδείξει ότι ο ορισμός αυτής της έννοιας είναι σύνθετος. Πολλές φορές συγχέονται με όρους όπως παραπληροφόρηση, ψευδείς ειδήσεις ή ακόμα και με τα «hoaxes» (τις απάτες). Πριν ορίσουμε τα fake news είναι χρήσιμο να αποδώσουμε την σημασία των όρων disinformation και misinformation, καθώς η ελληνική απόδοση και των δύο όρων είναι η παραπληροφόρηση, με αποτέλεσμα αρκετά συχνά να συγχέονται.

Ως disinformation ορίζονται οι ψεύτικες, ανακριβείς ή παραπλανητικές πληροφορίες, οι οποίες σχεδιάζονται παρουσιάζονται και προωθούνται με τέτοιο τρόπο, ώστε να προκαλέσουν σκόπιμα δημόσια ζημία ή να αποφέρουν κέρδος. Οι στόχοι αυτής της παραπληροφόρησης είναι δύο: αφενός τα οικονομικά οφέλη και αφετέρου οι πολιτικές ή ιδεολογικές σκοπιμότητες (European Comission, 2018). Ο όρος disinformation αποτελεί ένα πολύπλευρο πρόβλημα, αφού κάποιες μορφές παραπληροφόρησης σαν αυτή, έχουν δομηθεί λόγω της ανάπτυξης ορισμένων ψηφιακών μέσων. Το εν λόγω πρόβλημα αφορά πολιτικούς παράγοντες και μέσα ενημέρωσης. Αυτό συνεπάγεται το γεγονός ότι οι παράγοντες της πολιτικής μπορούν να επιδιώκουν σκοπίμως την παραπληροφόρηση και να την ενισχύουν. Ταυτόχρονα, τα ΜΜΕ μπορούν να διαδραματίσουν εξέχοντα ρόλο στην καταπολέμηση του φαινομένου της παραπληροφόρησης, αλλά αξίζει να σημειωθεί ότι δεν διατηρούν όλα τα μέσα τον ίδιο βαθμό ανεξαρτησίας.

Από την άλλη, όταν γίνεται λόγος για το misinformation εννοείται κάτι διαφορετικό από τον προηγούμενο όρο που αναλύθηκε, παρ' όλο που και οι δύο αναφέρονται ως παραπληροφόρηση. Ειδικότερα, ο όρος misinformation σημαίνει την εσφαλμένη πληροφόρηση, δηλαδή όταν μία παραπλανητική ή ανακριβής πληροφορία που διαδίδεται, δεν αναγνωρίζεται ως τέτοια.

Ύστερα, από τον ορισμό των όρων disinformation και misinformation, οφείλουμε να μετατοπιστούμε στην ανάλυση των fake news. Το Ινστιτούτο της Οξφόρδης για τις Μελέτες Υπολογισμού της Προπαγάνδας ορίζει τα fake news ως εξής: *«Παραπλανητικές, λανθασμένες πληροφορίες που θεωρούν ότι είναι αληθινά νέα για την πολιτική, την οικονομία ή τον πολιτισμό».* Επιπλέον, το λεξικό Cambridge αποδίδει στον όρο την παρακάτω σημασία: *«ψεύτικες (false) ιστορίες που εμφανίζονται ως ειδήσεις, διαδίδονται στο διαδίκτυο ή χρησιμοποιούν άλλα μέσα, συνήθως φτιαγμένα να επηρεάσουν πολιτικές θεωρήσεις/απόψεις ή ως αστεία».* Ακόμη, ο Jayson Harsin στο άρθρο του *«A critical guide to fake news: from comedy to tragedy»* (Harsin, 2018) υποστηρίζει ότι τα fake news είναι ειδήσεις που αναμειγνύουν την αλήθεια με το ψέμα για παραπλανητικούς σκοπούς ή ιστορίες που εφευρέθηκαν και δεν έχουν σχέση με την πραγματικότητα. Υποστηρίζεται, μάλιστα, η άποψη πως τα fake news αποτελούν μία μορφή προπαγάνδας, υπό την έννοια ότι το μήνυμα που μεταφέρουν έχει συγκεκριμένο στόχο. Αυτό παραπέμπει στον ορισμό που δίνει το Science Magazine για αυτές τις ειδήσεις. Πιο συγκεκριμένα, αναφέρεται ότι «*τα fake news είναι κατασκευασμένες πληροφορίες που μιμούνται το περιεχόμενο των ειδησεογραφικών μέσων σε μορφή, αλλά όχι οργανωτική διαδικασία ή πρόθεση*». Συχνά, άλλωστε, δημιουργείται η αίσθηση ότι

τα fake news είναι ένα είδος παραπληροφόρησης παρόμοιο με την προπαγάνδα που συνδέεται παραδοσιακά με το κράτος ή την κυβέρνηση. Παρ' όλα αυτά, είναι κομμάτι μιας ευρύτερης πολιτισμικής και ιστορικής μετατόπισης που περικλείεται στον όρο «μετα-αλήθεια» (post-truth).

Αναμφίβολα από τους παραπάνω ορισμούς προκύπτει ότι τα fake news επιτελούν συγκεκριμένους σκοπούς και περιγράφουν κάτι το οποίο δεν είναι αληθές. Ωστόσο, αυτός ο ορισμός τείνει περισσότερο προς τις πλαστές ειδήσεις και όχι τόσο προς τα fake news. Τα fake news διαφέρουν από αυτό, καθώς *«είναι γεγονότα που περιγράφονται σε ένα πλαίσιο ερμηνείας τους που αμφισβητεί την καθιερωμένη οπτική γωνία του κοινού περί πραγματικότητας, είτε έχουν συμβεί αυτά τα γεγονότα στο σύνολό τους είτε όχι. Συνεπώς τα fake news δεν είναι ψευδείς, αλλά ψευδεπίγραφες ειδήσεις»* (Πλειός, Fake News: ψευδεπίγραφες ειδήσεις. Εξαίρεση ή κανόνας ;, 2018)

Ο όρος fake news έχει διαδοθεί ιδιαίτερα τα τελευταία χρόνια και χρησιμοποιείται όλο και πιο συχνά με αποκορύφωμα το 2016, που χρησιμοποιήθηκε ευρύτατα στις προεδρικές αμερικανικές εκλογές. Έχει παρατηρηθεί ότι ο μέσος Αμερικανός συνάντησε μία έως τρεις ψευδεπίγραφες ειδήσεις κατά τη διάρκεια ενός μήνα πριν τις εκλογές στις Η.Π.Α. το 2016 (Science Mag.2018). Το 2019, σύμφωνα με μελέτη του Pew Research Center, οι Αμερικανοί χαρακτήριζαν το πρόβλημα των fake news ως μεγαλύτερο πρόβλημα κι από τον ρατσισμό και την κλιματική αλλαγή (Graham, 2019). Με την ανάπτυξη του ραδιοφώνου και της τηλεόρασης τον 20ο αιώνα, εξελίχθηκαν και οι σατιρικές ειδήσεις που μερικές φορές εκλαμβάνονταν ως αλήθεια.

Πέρα από τον τομέα της δημοσιογραφίας, ο όρος χρησιμοποιείται και στην πολιτική. Όμως, δεν πρόκειται για μία κατάσταση που διαδραματίζεται μόνο τα τελευταία χρόνια. Ήταν το 1835, όταν η νεοϋορκέζικη εφημερίδα *Sun* δημοσίευσε το λεγόμενο «Great Moon Hoax», δηλαδή έξι άρθρα σχετικά με την ανακάλυψη της ζωής στο φεγγάρι. Σύμφωνα με τον Jayson Harsin, τουλάχιστον από το 1999 οι ψευδεπίγραφες ειδήσεις χρησιμοποιήθηκαν σε μία ευρεία κλίμακα ως μέρος του αμερικανικού προγράμματος «The Daily Show» με τον Jon Stewart. Επρόκειτο για ένα ψεύτικο πρόγραμμα υπό την έννοια ότι μερικές φορές προσποιούνταν το ύφος των αληθινών ειδήσεων με δημοσιογράφους, οι οποίοι δημιουργούσαν ιστορίες ή σχολίαζαν στο στούντιο.

2.2. Χαρακτηριστικά και Κίνητρα

Αρχικά, όταν γίνεται αναφορά στα fake news ή αλλιώς τις ψευδεπίγραφες ειδήσεις, αυτό συνεπάγεται την ύπαρξη γεγονότων της επικαιρότητας με «ημερομηνία λήξης». Άρα, γίνεται λόγος για «νέα». Σύμφωνα με το Oxford English Dictionary η λέξη news (νέα) προέρχεται από την λατινική nova που σημαίνει νέα πράγματα. Στις αρχές του 18ου αιώνα τα «νέα» αναφέρονταν στις ειδήσεις των εφημερίδων, ενώ από την δεκαετία του 1920 αναφέρονται και σε ραδιοφωνικές εκπομπές και αρκετές δεκαετίες αργότερα στην τηλεόραση. Επιπλέον, αυτού του είδους οι ειδήσεις αναπαράγουν ένα μήνυμα, το οποίο έχει να κάνει με τα στερεότυπα και τις προκαταλήψεις, τα οποία είτε πρωτοεμφανίζονται την δεδομένη περίοδο ανάλυσης είτε είναι προϋπάρχοντα. Συγκεκριμένα, όταν παρουσιάζονται πληροφορίες που βρίσκονται μέσα στην δομή των πεποιθήσεών μας, επιβεβαιώνεται η ύπαρξη της προκατάληψης και δεχόμαστε ως αληθές το περιεχόμενο του μηνύματος. Με άλλα λόγια, οι ψευδεπίγραφες ειδήσεις σχετίζονται με τις προϋπάρχουσες πεποιθήσεις των ανθρώπων για τους πολιτικούς

ηγέτες, τα κόμματα, τους οργανισμούς και τα παραδοσιακά μέσα ενημέρωσης. Ορισμένες από αυτές είναι απλές κατασκευές, ενώ άλλες περιλαμβάνουν στοιχεία αλήθειας προκειμένου να φαίνονται αξιόπιστες. Η ψυχολογία έχει δείξει ότι πρωταρχικός παράγοντας για το αν οι άνθρωποι πιστεύουν ένα μήνυμα είναι εάν συμφωνούν ή όχι, καθώς η διαφωνία τους ενδέχεται να τους οδηγήσει στον έλεγχο της αξιοπιστίας του μηνύματος. Τα στερεότυπα που αναπαράγονται μέσω των ψευδεπίγραφων ειδήσεων αφορούν τόσο την ίδια την κοινωνία όσο και την πολιτική.

Παράλληλα, τα fake news προέρχονται από διαφορετικά σύνολα συμφερόντων, τα οποία αλληλεπικαλύπτονται: οικονομικοί στόχοι - πολιτικά αποτελέσματα, πολιτικοί στόχοι - οικονομικά αποτελέσματα. Ειδικότερα, η πολιτική εκμεταλλεύεται το επιχειρηματικό πλάνο των fake news, καθώς οι μεγάλες μηχανές αναζήτησης (π.χ. Google), τα μεγάλα δίκτυα διαφημίσεων (π.χ. Google ads) και τα Μέσα Κοινωνικής Δικτύωσης βοήθησαν τα fake news να προσελκύσουν το ενδιαφέρον και να τα συνενώσουν με τις εταιρείες διαφημίσεων. Οι θεωρίες συνωμοσίας, τα ψέματα και οι απάτες (hoaxes) διαδόθηκαν πιο αποτελεσματικά μέσω των Μέσων Κοινωνικής Δικτύωσης (ΜΚΔ), όπως το Facebook, το Twitter, το Snapchat κ.α. Βέβαια, υπάρχουν και πολιτικά κίνητρα για την παραγωγή ή την εκμετάλλευση των ψευδεπίγραφων ειδήσεων. Η επαγγελματική πολιτική επικοινωνία κατέληξε όλο και πιο συστηματικά να διαχειρίζεται τα μέσα ενημέρωσης και την κοινή γνώμη μέσω των fake news.

Όπως έχει ήδη προαναφερθεί, οι ψευδεπίγραφες ειδήσεις σχετίζονται και με τομείς σαν την πολιτική. Σύμφωνα με τον Γ. Πλειό (Πλειός, Fake News: Τα 4+1 βασικά γνωρίσματα, 2019),πρόκειται για όσα λέγονται από πολιτικούς και κυρίως εκείνα που λέγονται στην πολιτική διαδικασία. Οι ειδήσεις αυτές δεν κατασκευάζονται και αναπαράγονται μόνο από τους πολιτικούς. Σημαντικό ρόλο σε αυτό διακατέχουν τα μέσα ενημέρωσης, καθώς οι ειδήσεις που δημιουργούν ή αναπαράγουν αποτελούν ένα μέσο άσκησης κυριαρχίας μέσα από την ανακατασκευή της πραγματικότητας. Ωστόσο, όπως σημειώνει ο Γ. Πλειός στο προαναφερθέν άρθρο, «όσο περισσότερο τα μέσα ή οι δημοσιογράφοι έχουν ασθενή σχέση με κάποια ιδεολογία, ειδικά εκείνες που τις διαπερνούν οι αρχές της ισότητας και της κοινωνικής δικαιοσύνης, και αντιστοίχως έχουν πιο ισχυρή σχέση με την αγορά, τόσο πιθανότερο είναι να κατασκευάσουν ή να αναπαράγουν fake news σε μια αντιπαράθεση ή ανταγωνισμό».

Τα κίνητρα που κρύβονται πίσω από τις ειδήσεις τέτοιου περιεχομένου, αναμφίβολα, ποικίλλουν. Οι ψευδεπίγραφες ειδήσεις δεν αποτελούν τίποτε άλλο παρά ένα μέσο για την επίτευξη ενός στόχου. Οποιαδήποτε ανάρτηση θα μπορούσε να θεωρηθεί προκατειλημμένη ως έναν βαθμό, εκείνο, όμως, που διαφέρει στις «εκστρατείες των fake news» είναι ότι βασίζονται σε κατασκευασμένα γεγονότα, με σκοπό να τραβήξουν την προσοχή του αναγνώστη. Η απόκτηση χρημάτων ή δύναμης είναι σχεδόν πάντα παρούσες πίσω από τις ψευδεπίγραφες ειδήσεις (EAVI, 2017). Παρ' όλα αυτά μία δημοσίευση μπορεί να οφείλεται σε ιδεολογικά ή πολικά αίτια Σε αυτή την περίπτωση οι ψευδεπίγραφες ειδήσεις επιτελούν προπαγανδιστικό σκοπό. Ειδικότερα, αναπαράγονται από ορισμένες κυβερνήσεις, εταιρείες ή ακόμα και μη κερδοσκοπικούς οργανισμούς που στοχεύουν στον έλεγχο στάσεων, αξιών, συμπεριφορών και γνώσεων. Τα πολιτικά κίνητρα πίσω από τα fake news ενδέχεται να είναι η μεταβολή της γνώμης των ανθρώπων για διάφορες πολιτικές πεποιθήσεις ή κάποια άλλη γνώμη. Ιδιαίτερο χαρακτηριστικό είναι ότι αυτή η μορφή ειδήσεων στοχεύει ως επί το πλείστων στο συναίσθημα, προκειμένου να προσελκύσει τους αναγνώστες, οι οποίοι αντιμετωπίζονται ως στόχοι. Σε αυτή τη διαδικασία ιδιάζοντα ρόλο παίζουν οι

ψυχολογικές επιθυμίες των αναγνωστών, οι οποίες ικανοποιούνται μέσα από την επιβεβαίωση των προκαταλήψεών τους. Κατά συνέπεια, η παραπάνω διαδικασία μπορεί είτε να λειτουργήσει ευεργετικά, μέσω της προώθησης πολιτικών προσώπων ή ιδεολογιών, είτε να προκαλέσει βαρύτατες ζημίες σε πρόσωπα ή ομάδες μέσα από την διασπορά αναληθών πληροφοριών.

2.3. Fake News,Ψηφιακά Μέσα και Πολιτική

Με την έλευση του διαδικτύου στα τέλη του 20ου αιώνα και των μέσων κοινωνικής δικτύωσης στις αρχές του 21ου αιώνα πολλαπλασιάστηκε δραματικά ο κίνδυνος της παραπληροφόρησης, της προπαγάνδας και της εξαπάτησης. Τα παραδοσιακά μέσα ακολουθούν τους δημοσιογραφικούς κανόνες αντικειμενικότητας και ισορροπίας που προέκυψαν ως αντίδραση στην ευρεία χρήση προπαγάνδας κατά τον Α' Παγκόσμιο Πόλεμο. Ωστόσο, με την άφιξη του διαδικτύου και την χαμηλού κόστους είσοδο σε αυτό τον τομέα, πολλοί από τους κανόνες «θυσιάστηκαν» στο βωμό του ανταγωνισμού με τα παραδοσιακά μέσα και της υπονόμευσης τέτοιων επιχειρηματικών μοντέλων παραδοσιακών πηγών ειδήσεων που διέθεταν υψηλά επίπεδα εμπιστοσύνης και αξιοπιστίας στο κοινό. Το νέο περιβάλλον των μέσων ενημέρωσης είναι δυναμικό και συνεχίζει να αναπτύσσεται με νέους τρόπους και ρυθμούς έχοντας σοβαρές συνέπειες για την δημοκρατική διακυβέρνηση και τις πολιτικές πρακτικές. Η εικόνα μέσω των ΜΜΕ πλάθεται από την εμφύτευση και επανάληψη ιδεών προκειμένου να εντυπωθεί στο υποσυνείδητο. Η σχέση των πολιτών με την πολιτική, αλλά και τους πολιτικούς, έχει μεταβληθεί, καθώς έχει διευκολυνθεί η παραγωγή, η διάδοση και η ανταλλαγή πολιτικού περιεχομένου μέσα από πλατφόρμες και δίκτυα που χρησιμοποιούν την αλληλεπίδραση.

Πρώτα απ' όλα, αξίζει να παρατηρήσουμε τον ρόλο που διαδραματίζουν τα νέα ψηφιακά μέσα σε μια δημοκρατική κοινωνία. Θα μπορούσε κανείς να υποστηρίξει ότι πρωταρχικός τους ρόλος είναι να ενημερώνουν το κοινό παρέχοντάς του τις απαραίτητες πληροφορίες για την λήψη σημαντικών αποφάσεων σχετικών με την πολιτική, καθώς και να ελέγχουν τις κυβερνητικές κινήσεις. Έχουν την δυνατότητα, ως διαμορφωτές της πραγματικότητας, να ορίζουν την ατζέντα για δημόσια συζήτηση θεμάτων (agenda setting). Παράλληλα, παρέχουν ένα δημόσιο βήμα για την πολιτική έκφραση. Μάλιστα, τα τελευταία χρόνια έχει αποδειχτεί ότι τα μέσα κοινωνικής δικτύωσης είναι τα κύρια μέσα, στα οποία οι νέοι άνθρωποι επιλέγουν να αναπτύξουν τις πολιτικές τους ταυτότητες (Wooley & Howard, 2017). Παρ' όλα αυτά, τα ΜΚΔ μπορεί να αποτελούν σημαντική πηγή πολιτικών ειδήσεων – ιδιαίτερα σε μια προεκλογική περίοδο – ωστόσο δεν είναι η κυρίαρχη πηγή τέτοιων ειδήσεων, όπως έγινε γνωστό για τις προεδρικές εκλογές των Η.Π.Α. το 2016 (Allcott & Gentzkow, 2017).

Ακριβώς επειδή η πρόσβαση σε αυτά γίνεται χωρίς ιδιαίτερα εμπόδια, οι πολιτικοί στράφηκαν στα νέα μέσα για να παρακάμψουν τον έλεγχο του «διαμεσολαβητή», δηλαδή των παραδοσιακών μέσων, στην ατζέντα των ειδήσεων. Με αυτό τον τρόπο, οι πολιτικοί μπορούν να κοινοποιήσουν τις πολιτικές τους απόψεις σε ένα ευρύ κοινό (followers) χωρίς το «φιλτράρισμα» των παραδοσιακών μέσων, καλλιεργώντας ταυτόχρονα μία αμεσότερη σχέση με αυτό, αλλά κι ένα προφίλ πιο «ανθρώπινο». Όπως επισημαίνει ο κος Σ. Παπαθανασόπουλος σε κείμενό του στην Ελληνική Εταιρεία Πολιτικής Επιστήμης, «*έχει αναδυθεί η κουλτούρα των μέσων (media culture) στο πλαίσιο της οποίας εικόνες, ήχοι και θεάματα υποβοηθούν στην παραγωγή του ιστού της*

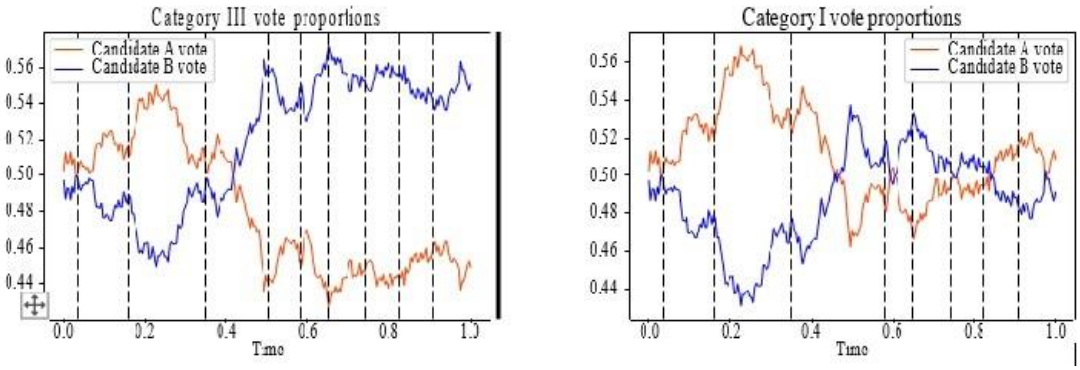
καθημερινής ζωής, στην κυριαρχία του ελεύθερου χρόνου, τη διαμόρφωση των πολιτικών απόψεων και την κοινωνική συμπεριφορά, καθώς και στην παροχή υλικού μέσα από το οποίο οι άνθρωποι διαμορφώνουν τις ταυτότητές τους» (Παπαθανασόπουλος, 2017).

Είναι χαρακτηριστική η περίπτωση του σκανδάλου Cambridge Analytica για τον τρόπο που προσπάθησαν οι πολιτικοί να εκμεταλλευτούν τα ΜΚΔ. Το Δεκέμβριο του 2015 η εφημερίδα Guardian αποκάλυψε την παράνομη χρήση δεδομένων των χρηστών του Facebook από την εταιρεία πολιτικών αναλύσεων Cambridge Analytica.Όπως αναφέρθηκε και προηγουμένως η πλειοψηφία του κόσμου πιστεύει την είδηση που θέλει να πιστέψει, δηλαδή την είδηση που επιβεβαιώνει τα στερεότυπα του και τις προκαταλήψεις του. Έτσι, λοιπόν, σκοπός της χρήσης αυτής ήταν η δημιουργία προφίλ των χρηστών, έτσι ώστε τα διαδιδόμενα Fake News να είναι πιο εξατομικευμένα. (Παπαϊωάννου, 2018) Για την υπόθεση αυτή επιβλήθηκε στο Facebook πρόστιμο 700.000 ευρώ από το βρετανικό Γραφείο Επιτροπείας των Πληροφοριών για διαρροή προσωπικών δεδομένων χρηστών που χρησιμοποιήθηκαν για πολιτική διαφήμιση και προβολή υποψηφίων στις ΗΠΑ και στη Βρετανία. Για την ίδια υπόθεση το Facebook πλήρωσε πρόστιμο ύψους πέντε εκατομμυρίων δολαρίων στις ΗΠΑ (Εφημερίδα των Συντακτών, 2019).Η επίδραση, όμως, των Fake News στην πολιτική γίνεται περισσότερο κατανοητή με την αξιοποίηση των γραφημάτων που παρουσιάζονται στην συνέχεια (Εικ. 1, Εικ. 2)

Σε σχετική μελέτη ερευνητές απέδειξαν, μέσω ενός μαθηματικού μοντέλου, τον ισχυρισμό ότι τα Fake News έχουν σημαντικές συνέπειες στην πολιτική ζωή (Brody & Meier, 2018). Αξιοποιώντας βασικές αρχές από την θεωρία της επικοινωνίας για τη ροή της πληροφορίας, μοντελοποίησαν τα fake news ως έναν παράγοντα «θορύβου» στο σήμα της πληροφορίας που διαδίδεται υποθέτοντας για χάρη απλότητας ότι το σήμα αυτό θα έχει μια γραμμική μορφή. Συγκεκριμένα η ανάλυση οδήγησε στην μαθηματική σχέση:

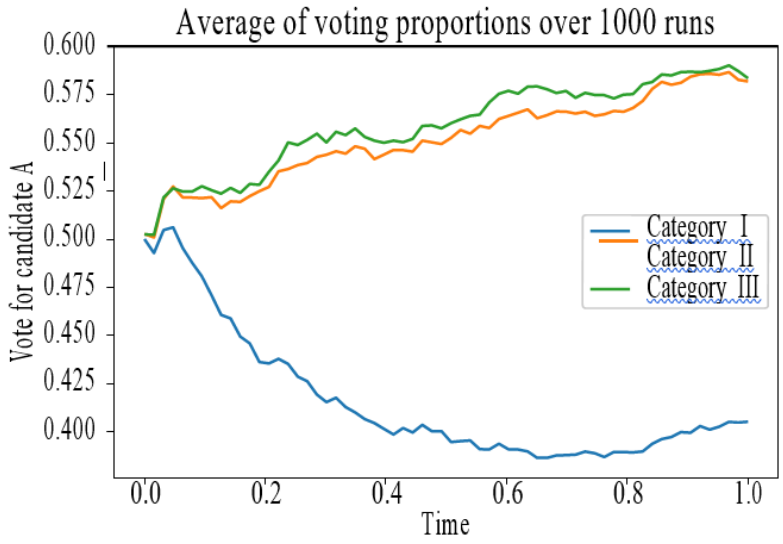
$$\eta_t=\sigma X_t+B_t+F_t$$

,όπου η_t η ροή της πληροφορίας, X_t η αξιόπιστη γνώση, σ ο σταθερός ρυθμός διάδοσης της αξιόπιστης γνώσης , B_t ένας παράγοντας «θορύβου» ή αβεβαιότητας και F_t τα Fake News.Στη συνέχεια δημιούργησαν την προσομοίωση μιας εκλογικής αναμέτρησης ανάμεσα σε δυο υποψηφίους Α και Β, υποθέτοντας τρεις κατηγορίες ψηφοφόρων. Οι ψηφοφόροι της κατηγορίας 1 είναι αυτοί που δεν έχουν καν επίγνωση ότι μπορεί μέρος της ενημέρωσης τους να είναι Fake News. Οι ψηφοφόροι της κατηγορίας 2 είναι αυτοί που είναι ενήμεροι για την ύπαρξη πιθανής παραπληροφόρησης μέσα στην ενημέρωση τους αλλά δεν γνωρίζουν ακριβώς τη στιγμή που λαμβάνουν αυτά τα μηνύματα. Τέλος, οι ψηφοφόροι της κατηγορίας 3 είναι αυτοί που μπορούν να ξεχωρίζουν αμέσως την ψευδή είδηση και να την αποκόπτουν από την ενημέρωση τους. Τα αποτελέσματα της προσομοίωσης αυτής βλέπουμε στην παρακάτω εικόνα :



Εικόνα 1:Πηγή : «How to Model Fake News» ,Dorje C. Brody and David M. Meier,2018

Στην εικόνα 1 βλέπουμε αριστερά μια εύκολη επικράτηση του υποψήφιου Β υποθέτοντας ότι οι ψηφοφόροι ήταν της κατηγορίας 3, δηλαδή ότι τα Fake News δεν παίζουν κανένα ρόλο σε αυτό το αποτέλεσμα. Αντιθέτως, παρατηρούμε στο δεξί σκέλος, υποθέτοντας ότι οι ψηφοφόροι είναι της κατηγορίας 1, εκθέτοντας τους σε Fake News σε 9 περιπτώσεις (οι κατακόρυφες γραμμές), τελικά ο υποψήφιος Α καταφέρνει να πάρει τη νίκη έστω και για μικρή διαφορά. Ας δούμε όμως και άλλη μια προσομοίωση από την ίδια έρευνα (εικ.2).



Εικόνα 2:«How to Model Fake News», Dorje C. Brody and David M. Meier, 2018

Στο γράφημα αυτό βλέπουμε και πάλι την προσομοίωση μιας εκλογικής αναμέτρησης και συγκεκριμένα το ποσοστό του πληθυσμού που ψηφίζει τον υποψήφιο Α σε συνάρτηση με το χρόνο. Παρατηρούμε εύκολα τη μεγάλη απόκλιση που έχουν οι ψηφοφόροι της κατηγορίας 1 στον τρόπο που ψηφίζουν σε σχέση με τις άλλες δύο κατηγορίες. Από το σύνολο των γραφημάτων αυτό που προκύπτει ξεκάθαρα είναι η επίδραση των fake news στην πρόθεση ψήφου αλλά ακόμα πιο έντονο είναι το γεγονός ότι η επίδραση τους είναι συντριπτικά μεγαλύτερη στα κομμάτια του πληθυσμού που δεν γνωρίζουν καν την ύπαρξη των fake news και άρα δεν μπορούν να προστατευτούν από αυτά.

Από τα παραπάνω, μπορεί εύκολα να συμπεράνει ο αναγνώστης ότι στην εποχή της πληροφορίας τα μέσα ενημέρωσης και κοινωνικής δικτύωσης διαδραματίζουν ίσως τον μεγαλύτερο ρόλο στην διαμόρφωση απόψεων (πολιτικών κ.α.). Αυτό αποτελεί μία σημαντική παράμετρο, την οποία οι εκούσιοι παραγωγοί ψευδεπίγραφων ειδήσεων λαμβάνουν σοβαρά υπ' όψιν, όταν στοχεύουν στην εξυπηρέτηση πολιτικών σκοπών. Δεν είναι τυχαίο, άλλωστε, που τα ΜΚΔ χρησιμοποιούνται ενεργά για την χειραγώγηση της κοινής γνώμης. Σύμφωνα με τον Αμερικανό συγγραφέα Ralph Keyes, η κοινωνία έχει επέλθει σε μία εποχή μετά-αλήθειας. Η εξαπάτηση είναι το κυρίαρχο χαρακτηριστικό της σύγχρονης ζωής, ενώ η ευρεία εξάπλωσή του έχει οδηγήσει στην απευαισθητοποίηση των ανθρώπων.

2.4. Επιρροή στην Επιστήμη

Αναλογιζόμενοι την επίδραση των Fake News στην πολιτική συμπεραίνουμε ότι η αντιμετώπιση του φαινομένου αφορά όλους τους πολίτες, με την έννοια, ότι απειλείται η ίδια η δημοκρατία, τότε μπορεί να αναλογιστεί κανείς τις επικίνδυνες συνέπειες που μπορεί να έχει η εξάπλωση των Fake News σε επιστημονικά ζητήματα.

Τα ζητήματα που εμφανίζονται στον τομέα της Ιατρικής και της Βιολογίας αποτελούν δύο από τους σημαντικότερους κλάδους αναφοράς . Αν σε ατομικό επίπεδο δεν έχει ο κάθε πολίτης πρόσβαση σε αξιόπιστη γνώση σε ζητήματα που αφορούν την υγεία τότε καταλαβαίνουμε πολύ εύκολα ότι δημιουργείται κίνδυνος για τον ίδιο αλλά και για τη δημοσία υγεία. Ένα ξεκάθαρο παράδειγμα είναι η παραπληροφόρηση που διαδόθηκε πριν από κάποια χρόνια κυρίως μέσω ΜΚΔ γύρω από τα εμβόλια για τις παιδικές ασθένειες της ιλαράς, της παρωτίτιδας και της ερυθράς ότι προκαλούν αυτισμό (Rao & Andrade, 2011), (Hviid, Hansen, Frisch, & Melbye, 2019). Αυτό είχε ως αποτέλεσμα, όχι μόνο να αυξηθούν οι μολύνσεις από ιλαρά στην Ευρώπη (WHO, 2018) αλλά και να ενισχυθεί το ρεύμα υπέρ της καχυποψίας για τα εμβόλια (Kestenbaum & Feemster, 2015), (Larson, 2018).

Βεβαίως, τα παραδείγματα είναι πολλά και επεκτείνονται και πέρα από τους τομείς υγείας. Φαινόμενο των τελευταίων χρόνων είναι η άρνηση της κλιματικής αλλαγής, την ίδια στιγμή που επιστήμονες σε όλον τον κόσμο επισημαίνουν τις αρνητικές συνέπειες που μπορεί να έχει στον πλανήτη τα επόμενα χρόνια το φαινόμενο (Watts, 2018). Άξια αναφοράς, ωστόσο, είναι και η ύπαρξη των λεγόμενων «flat earthers» δηλαδή ανθρώπων που πιστεύουν ότι η γη είναι επίπεδη, των οποίων η ιδεολογία μέσω της ταχύτητας και ευκολίας της διασποράς των fake news αποκτά όλο και μεγαλύτερο αριθμό υποστηρικτών.

Αρκετοί τέτοιου είδους ισχυρισμοί μπορούν πολύ εύκολα να αποδομηθούν με μια απλή αναζήτηση στο Διαδίκτυο. Ωστόσο, αυτό που ξεχωρίζει τον επιστημονικό ισχυρισμό από ψευδο-επιστημονικούς ισχυρισμούς είναι ότι ο πρώτος σε κάθε περίπτωση μπορεί να ελεγχθεί γιατί έχει βασιστεί σε συγκεκριμένη επιστημονική μεθοδολογία. Αντίθετα ένας ψευδο-επιστημονικός ισχυρισμός ακριβώς επειδή μπορεί να μην βασίζεται στην αποδεικτική μέθοδο, συχνά είναι και δύσκολο να αποδομηθεί και για αυτό βρίσκει χώρο στο δημόσιο λόγο με φαινομενικά αδιάσειστα επιχειρήματα (Stemwedel, 2011).

Οι λόγοι και τα κίνητρα, επίσης, των ατόμων που ελλοχεύουν πίσω από τέτοιους ισχυρισμούς ποικίλλουν. Άλλοτε, τα κίνητρα είναι προσωπική προβολή ενώ άλλοτε μπορεί να είναι εξυπηρέτηση συγκεκριμένων πολιτικών ή οικονομικών συμφερόντων, όπως για παράδειγμα η συστηματική απόκρυψη ή και αλλοίωση στοιχείων από

κυβερνήσεις και βιομηχανίες καπνού, που αφορούν τις πιθανές βλάβες από το κάπνισμα (Brownell & Warner, 2009), (Smith, Thompson, & Lee, 2016).

Έτσι, λοιπόν, ο τρόπος για να υποστηριχθούν τέτοιες απόψεις είναι, κατά διαστήματα, να επανέρχονται στο προσκήνιο, εκμεταλλευόμενοι ίσως την επικαιρότητα και να αναπαράγουν Fake News για να διαδώσουν την άποψη τους βασιζόμενη σε σύγχρονες ειδήσεις.

Σε κάθε περίπτωση τέτοιου είδους διασπορά Fake News έχει κλονίσει τη σχέση κοινωνίας και επιστήμης με μεγάλος μέρος του πληθυσμού να έχει χάσει την εμπιστοσύνη του σε αυτήν (Funk, 2017),επιφέροντας όλες τις σχετικές επιπτώσεις.

ΚΕΦΑΛΑΙΟ 3
ΝΕΕΣ ΤΕΧΝΟΛΟΓΙΕΣ ΚΑΙ ΔΙΑΔΟΣΗ ΤΩΝ FAKE NEWS

3.1. Τεχνητή Νοημοσύνη

3.1.1. Ορισμός και Εφαρμογές

Η Τεχνητή Νοημοσύνη (Artificial Intelligence - AI) αναφέρεται στην ικανότητα μιας μηχανής να αναπαράγει τις γνωστικές λειτουργίες ενός ανθρώπου, όπως είναι η μάθηση, ο σχεδιασμός και η δημιουργικότητα. Η τεχνητή νοημοσύνη καθιστά τις μηχανές ικανές να «κατανοούν» το περιβάλλον τους, να επιλύουν προβλήματα και να δρουν προς την επίτευξη ενός συγκεκριμένου στόχου. Ο υπολογιστής λαμβάνει δεδομένα (ήδη έτοιμα ή συλλεγμένα μέσω αισθητήρων, π.χ. κάμερας), τα επεξεργάζεται και ανταποκρίνεται βάσει αυτών. Τα συστήματα τεχνητής νοημοσύνης είναι ικανά να προσαρμόζουν τη συμπεριφορά τους, σε ένα ορισμένο βαθμό, αναλύοντας τις συνέπειες προηγούμενων δράσεων και επιλύοντας προβλήματα με αυτονομία (Ευρωπαϊκό Κοινοβούλιο, 2021) .

Η Τεχνητή Νοημοσύνη βρίσκει εφαρμογή σε διαφορετικά επιστημονικά πεδία που εκτείνονται από τη Φυσική, τη Βιολογία και την Πληροφορική μέχρι και τις θεωρητικές επιστήμες όπως η Κοινωνιολογία και η Γλωσσολογία. Βασικές εφαρμογές της Τεχνητής Νοημοσύνης που αντανακλούν και στην καθημερινότητα μας είναι (Ευρωπαϊκό Κοινοβούλιο, 2021):

- Διαδικτυακές αγορές και διαφήμιση: Η τεχνητή νοημοσύνη χρησιμοποιείται ευρέως για την παροχή εξατομικευμένων συστάσεων, για παράδειγμα βάσει προηγούμενων αναζητήσεων και αγορών ή άλλων συμπεριφορών. Παίζει, επίσης, εξαιρετικά σημαντικό ρόλο στον κλάδο του εμπορίου, καθώς χρησιμοποιείται για τη βελτιστοποίηση προϊόντων, τον προγραμματισμό των αποθεμάτων, τον εφοδιαστικό τομέα.
- Διαδικτυακή αναζήτηση: Οι μηχανές αναζήτησης παρέχουν αποτελέσματα βάσει της τεράστιας ποσότητας δεδομένων που εισάγουν οι χρήστες στο διαδίκτυο.
- Προσωπικοί ψηφιακοί βοηθοί: Τα έξυπνα τηλέφωνα (smartphones) χρησιμοποιούν την τεχνητή νοημοσύνη για την παροχή βελτιστοποιημένων και εξατομικευμένων ρυθμίσεων στους χρήστες τους. Ο εικονικός βοηθός λειτουργεί ως προσωπικός γραμματέας του χρήστη: απαντά σε ερωτήσεις, παρέχει συστάσεις, υπενθυμίζει συναντήσεις. Είναι επίσης ένας ηλεκτρονικός συνομιλητής που προσαρμόζεται στα ατομικά χαρακτηριστικά ενός συγκεκριμένου ατόμου, λαμβάνοντας υπόψη το περιβάλλον του χρήστη, το εύρος των ενδιαφερόντων του και τις συνήθειες του.
- Αυτόματες μεταφράσεις: Τα λογισμικά αυτόματης μετάφρασης και υποτιτλισμού, που βασίζονται είτε σε γραπτό είτε σε προφορικό λόγο,

χρησιμοποιούν τη τεχνητή νοημοσύνη για την παροχή και βελτίωση μεταφράσεων.

- Έξυπνα σπίτια, πόλεις και υποδομές: Οι έξυπνοι θερμοστάτες αναλύουν τη συμπεριφορά μας προκειμένου να αποθηκεύσουν ενέργεια, ενώ οι έξυπνες πόλεις βασίζονται σε ευφυή συστήματα ρύθμισης της κυκλοφορίας για να βελτιώσουν τη συνδεσιμότητα και να μειώσουν την κυκλοφοριακή συμφόρηση.
- Αυτοκίνητα: Παρότι τα αυτόνομα οχήματα δεν αποτελούν ακόμα μέρος της καθημερινότητάς μας, τα αυτοκίνητα απαρτίζονται ήδη από ευφυή συστήματα ασφαλείας που κάνουν χρήση τεχνητής νοημοσύνης. Η ΕΕ, για παράδειγμα, συμμετείχε στη χρηματοδότηση των αυτόματων αισθητήρων VI-DAS που εντοπίζουν ενδεχόμενες καταστάσεις κινδύνου και ατυχήματα.
- Κυβερνοασφάλεια: Τα συστήματα τεχνητής νοημοσύνης μπορούν να συμβάλουν στην αναγνώριση και αντιμετώπιση επιθέσεων και απειλών στον κυβερνοχώρο βάσει της συνεχόμενης εισροής δεδομένων.
- Χρήσεις κατά του COVID-19: Στην περίπτωση του COVID-19, η τεχνητή νοημοσύνη έχει χρησιμοποιηθεί σε συσκευές θερμικής απεικόνισης σε αεροδρόμια και αλλού. Στην ιατρική, η ΤΝ μπορεί να συμβάλει στην αποτελεσματική διάγνωση του κορονοϊού μέσω της χρήσης αλγορίθμων που μελετούν υπολογιστικές τομογραφίες θώρακα. Μπορεί, επίσης, να βοηθήσει στην παρακολούθηση της εξάπλωσης του ιού μέσω της παροχή δεδομένων.
- Καταπολέμηση της παραπληροφόρησης: Ορισμένες εφαρμογές τεχνητής νοημοσύνης μπορούν να συμβάλουν στην ανίχνευση των ψευδών ειδήσεων και της παραπληροφόρησης στα κοινωνικά δίκτυα μέσω του εντοπισμού συγκεκριμένων λέξεων και εκφράσεων αλλά και αξιόπιστων πηγών πληροφόρησης.

Τα προαναφερθέντα παραδείγματα αποτελούν μερικούς από τους κύριους τομείς χρήσης και αξιοποίησης της τεχνολογίας της τεχνητής νοημοσύνης. Όλες, όμως, αυτές οι λειτουργίες επιτυγχάνονται χάρη στην πιο ενδιαφέρουσα διαδικασία που αφορά το κομμάτι της τεχνητής νοημοσύνης και αυτή που έχουν επικεντρώσει την έρευνα τους πάρα πολλοί επιστήμονες, τη διαδικασία της μάθησης.

3.1.2. Μηχανική μάθηση

Η Μηχανική Μάθηση είναι υποπεδίο της επιστήμης των υπολογιστών που αναπτύχθηκε από τη μελέτη της αναγνώρισης προτύπων και της υπολογιστικής θεωρίας μάθησης στην τεχνητή νοημοσύνη. Ο ορισμός της μηχανικής μάθησης δόθηκε το 1959, από τον Άρθουρ Σάμουελ, ορίζοντάς την ως «Πεδίο μελέτης που δίνει στους υπολογιστές την ικανότητα να μαθαίνουν, χωρίς να έχουν ρητά προγραμματιστεί». Η μηχανική μάθηση, αν και έχει λάβει ποικίλες μορφές κατά την πορεία της εξελικτικής της πορείας, οι βασικές της λειτουργίες αποτελούν την διερεύνηση της μελέτης και της

κατασκευής αλγορίθμων, ικανών να εκπαιδεύονται και να αυξάνουν την αρχική τους γνώση βασιζόμενοι σε δεδομένα που τους έχουν δοθεί ως εναρκτήρια στοιχεία, και στην συνέχεια να δημιουργούν προβλέψεις επί αυτών για μελλοντικές καταστάσεις. Τέτοιοι αλγόριθμοι, επομένως, λειτουργούν κατασκευάζοντας μοντέλα από πειραματικά δεδομένα, προκειμένου οι προβλέψεις βασιζόμενες σε αυτά να είναι ακριβείς ή να εξάγουν αποφάσεις που εκφράζονται ως το αποτέλεσμα (Wikipedia, Μηχανική Μάθηση).

Βάσει του ορισμού αυτού, η Μηχανική Μάθηση έχει ως σκοπό τη δημιουργία μηχανών ικανών να μαθαίνουν, να βελτιώνουν την απόδοσή τους σε κάποιους τομείς, μέσω της αξιοποίησης προηγούμενης γνώσης και εμπειρίας. Ένας αντίστοιχος γενικός ορισμός Μηχανικής Μάθησης δίνεται από τον Mitchell (1997):

> *«Ένα πρόγραμμα υπολογιστή λέμε ότι μαθαίνει από την εμπειρία E ως προς κάποια κλάση εργασιών T και μέτρο απόδοσης P, αν η απόδοσή του σε εργασίες από το T, όπως μετριέται από το P, βελτιώνεται μέσω της εμπειρίας E.»*

Η μηχανική μάθηση είναι στενά συνδεδεμένη και συχνά συγχέεται με την υπολογιστική στατιστική. Ο κλάδος αυτός επικεντρώνεται στην πρόβλεψη καταστάσεων και συμπεριφορών μέσω της χρήσης των υπολογιστών. Έχει ισχυρούς δεσμούς με την μαθηματική βελτιστοποίηση, η οποία παρέχει τις μεθόδους, τη θεωρία και τους τομείς εφαρμογής. Η Μηχανική Μάθηση εφαρμόζεται σε μια σειρά από υπολογιστικές εργασίες, οι οποίες χαρακτηρίζονται από ακρίβεια και αυστηρή ανάλυση. Τόσο ο σχεδιασμός τους όσο και ο ρητός προγραμματισμός των αλγορίθμων είναι ανέφικτος. Παραδείγματα εφαρμογών αποτελούν τα φίλτρα spam (spam filtering), η οπτική αναγνώριση χαρακτήρων (OCR), οι μηχανές αναζήτησης και η υπολογιστική όραση. Η μηχανική μάθηση μερικές φορές συγχέεται με την εξόρυξη δεδομένων, της οποίας ο σκοπός διαφέρει από αυτόν της μηχανικής μάθησης. Η τελευταία επικεντρώνεται περισσότερο στην εξερευνητική ανάλυση των δεδομένων, γνωστή και ως μη επιτηρούμενη μάθηση (Wikipedia, Μηχανική Μάθηση).

Επιπροσθέτως, η μηχανική μάθηση μπορεί να εφαρμοστεί στο πεδίο της ανάλυσης δεδομένων όπου χρησιμοποιείται για την επινόηση πολύπλοκων μοντέλων και αλγορίθμων που οδηγούν στην πρόβλεψη. Τα αναλυτικά μοντέλα επιτρέπουν στους ερευνητές, τους επιστήμονες δεδομένων, τους μηχανικούς και τους αναλυτές να παράγουν αξιόπιστες αποφάσεις και αποτελέσματα και να αναδείξουν αλληλο-συσχετίσεις μέσω της μάθησης από ιστορικές σχέσεις και τάσεις στα δεδομένα. Αξίζει όμως η αναζήτηση του λόγου για τον οποίο η μηχανική μάθηση βρήκε έντονη εφαρμογή τα τελευταία χρόνια.

Ο πρώτος και βασικότερος λόγος είναι ότι η μηχανική μάθηση για να παράξει αποτελεσματικές προβλέψεις χρειάζεται μεγάλο όγκο δεδομένων. Με τις σημερινές τεχνολογίες το πρόβλημα αυτό ξεπεράστηκε καθώς πλέον είναι πολύ πιο εύκολη και η συλλογή αλλά και η αποθήκευση μεγάλου όγκου δεδομένων. Καταλυτικό ρόλο στην αποτελεσματικότητα της Μηχανικής Μάθησης έπαιξε επίσης η μεγάλη τεχνολογική εξέλιξη στο hardware και συγκεκριμένα στις κάρτες γραφικών (GPU). Οι νέες GPU μπορούν και μας παρέχουν την υπολογιστική δύναμη που καθιστά δυνατή την ανάλυση και επεξεργασία αυτού του όγκου δεδομένων. Τέλος η ανάπτυξη της Υπολογιστικής Νέφους (Cloud Computing), μείωσε σημαντικά το κόστος υπηρεσιών απαραίτητων για την απόδοση ενός υπολογιστή, δημιούργησε δυνατότητες απεριόριστης αποθήκευσης

και παράλληλα έκανε ευκολότερη την πρόσβαση στα δεδομένα, με αποτέλεσμα σημαντικές διεργασίες γύρω από τη Μηχανική Μάθηση να έχουν πλέον διευκολυνθεί και επιταχυνθεί (Xenopoulos, 2017).

Η επίλυση ενός προβλήματος Μηχανικής Μάθησης απαιτεί μια συγκεκριμένη διαδικασία, προκειμένου να επιτευχθεί το βέλτιστο δυνατό αποτέλεσμα και να αποφευχθούν πιθανά λάθη. Η διαδικασία αυτή αφορά την κατανόηση του προβλήματος και των ιδιαίτερων χαρακτηριστικών του, τη συλλογή και επεξεργασία των δεδομένων αλλά και την επιλογή του κατάλληλου αλγορίθμου για το πρόβλημα μας. Μπορούμε, λοιπόν, να εξειδικεύσουμε την παραπάνω διαδικασία στα εξής βήματα (Yufeng, 2017):

1. Κατανόηση του προβλήματος: Όπως σε κάθε πρόβλημα, έτσι και εδώ ξεκινώντας την επίλυση είναι απαραίτητη η πλήρης κατανόηση του προβλήματος, τα ζητούμενα του, καθώς και ποιοτικά στοιχεία του, όπως είναι η αναγνώριση εξαρτημένων και ανεξάρτητων μεταβλητών.

2. Συλλογή Δεδομένων: Είναι ίσως το πιο σημαντικό βήμα μιας και η επίδοση ενός μοντέλου Μηχανικής Μάθησης εξαρτάται κυρίως από τα δεδομένα. Θα πρέπει, λοιπόν, τα δεδομένα να είναι αρκετά, να είναι αξιόπιστα και να είναι αντιπροσωπευτικά.

3. Επεξεργασία των Δεδομένων: Πριν ένα μοντέλο εφαρμοστεί, θα πρέπει να είμαστε σίγουροι ότι τα δεδομένα μας πληρούν τις παραπάνω προϋποθέσεις. Οπότε ,σε αυτό το βήμα ερχόμαστε αντιμέτωποι με διάφορα προβλήματα που μπορεί να εμφανιστούν, όπως χαμένα δεδομένα ή ασυνεπή δεδομένα αλλά ακόμα και με ενέργειες οι οποίες θα κάνουν τη ζωή μας πιο εύκολη στην κατανόηση των δεδομένων και στον χειρισμό τους, όπως είναι η κανονικοποίηση τους, η προσθήκη ετικετών ή η αποκοπή των άχρηστων δεδομένων για το πρόβλημα μας.

4. Εξερεύνηση των Δεδομένων: Είναι το βήμα στο οποίο, μαθαίνουμε καλύτερα τον «εχθρό» μας. Μέσω της εμφάνισης των δεδομένων ή της οπτικοποίησης τους σε γραφικές παραστάσεις , μπορούμε να δούμε χαρακτηριστικά των δεδομένων μας όπως είναι η γραμμικότητα, η επαναληψιμότητα, η μοναδικότητα ,τα οποία πιθανώς να μας βοηθήσουν να επιλέξουμε τον αλγόριθμο Μηχανικής Μάθησης που θα εφαρμόσουμε στο πρόβλημα μας.

5. Εφαρμογή Μαθηματικού Μοντέλου: Σε αυτό το σημείο επιλέγουμε τον κατάλληλο αλγόριθμο που θα εφαρμόσουμε στα δεδομένα μας. Ο τομέας της Μηχανικής Μάθησης ανάλογα με τον τρόπο μάθησης που εμπίπτει στη φύση του εκάστοτε προβλήματος αναπτύσσει τρία είδη μάθησης (Ευτυχίου, 2019):

    i. Επιβλεπόμενη Μάθηση (Supervised learning): ο αλγόριθμος δημιουργεί μια συνάρτηση λαμβάνοντας ως εισόδους ένα σύνολο στιγμιότυπών εκπαίδευσης με γνωστές εξόδους. Στόχος της είναι η γενίκευση της συνάρτησης ώστε να απεικονίζονται και δεδομένα εισόδου με άγνωστη έξοδο. Εφαρμογή βρίσκει σε προβλήματα ταξινόμησης (classification), πρόγνωσης (prediction) και διερμηνείας (Interpretation).

ii. Μη επιβλεπόμενη Μάθηση (Unsupervised Learning): ο αλγόριθμος προσπαθεί να ανακαλύψει τυχόν συσχετίσεις μεταξύ των στιγμιότυπών εισόδου με άγνωστη έξοδο προκειμένου να βρεθούν δομικοί σχηματισμοί τους. Εφαρμογή βρίσκει σε προβλήματα ανάλυσης συσχετισμών (association analysis) και ομαδοποίησης (clustering).

iii. Ενισχυτική Μάθηση (Reinforcement Learning): ο αλγόριθμος μαθαίνει μια στρατηγική ενεργειών μέσα από την άμεση παρατήρηση του τρόπου λειτουργίας ενός δυναμικού περιβάλλοντος, χωρίς κάποιος να του υποδείξει το υλικό εκπαίδευσης. Εφαρμογή βρίσκει σε προβλήματα σχεδιασμού (planning) όπως είναι ένα παιχνίδι σκάκι ή κίνηση ενός ρομπότ.

Λαμβάνοντας, λοιπόν, υπόψη τα παραπάνω αλλά και πιθανώς τα ειδικά γνωρίσματα των δεδομένων και του προβλήματος μας (π.χ αριθμητικά ή κατηγορικά δεδομένα) επιλέγουμε το κατάλληλο μοντέλο και το «εκπαιδεύουμε» με τα δεδομένα μας.

6. Αξιολόγηση Μοντέλου: Γενικά χρησιμοποιούνται διαφορετικές μετρικές ποσοτικής ή / και ποιοτικής αξιολόγησης για να αποδειχθεί η αποτελεσματικότητα του εκπαιδευμένου μοντέλου. Παραδείγματα αποτελούν η απόδοση σε πραγματικό χρόνο του μοντέλου, αποτύπωμα μνήμης του μοντέλου, ακρίβεια, το ψευδοθετικό και ψευδοαρνητικό ποσοστό, η λογαριθμική απώλεια κ.ά.

3.2. Επεξεργασία Φυσικής Γλώσσας

3.2.1. Ορισμός και Εφαρμογές

Η επεξεργασία φυσικής γλώσσας ευρέως διαδεδομένη με τον αγγλικό όρο Natural Language Processing (NLP) αναφέρεται στον κλάδο της επιστήμης υπολογιστών και πιο συγκεκριμένα στον κλάδο της τεχνητής νοημοσύνης που ασχολείται με τεχνολογίες οι οποίες βοηθούν τον υπολογιστή να αποκτήσει την ικανότητα να καταλαβαίνει κείμενο και ομιλίες φυσικής γλώσσας με τον ίδιο τρόπο που μπορούν να το κάνουν και οι άνθρωποι (Garbade, 2018).

Η επεξεργασία φυσικής γλώσσας στην ουσία κρύβεται πίσω από πολύ συνηθισμένες και γνωστές μας εφαρμογές μερικές από τις οποίες αναφέρθηκαν και νωρίτερα όπως (Garbade, 2018):

- Εφαρμογές Αυτόματης Μετάφρασης (π.χ. Google Translate)
- Αυτόματη Διόρθωση γραμματικών λαθών σε εφαρμογές (π.χ.Microsoft Word)
- Η Διαδραστική Φωνητική Απόκριση που χρησιμοποιείται σε τηλεφωνικά κέντρα για να απαντήσουν σε συγκεκριμένα αιτήματα χρηστών
- Εφαρμογές ψηφιακού προσωπικού βοηθού (π.χ Google Assistance, Siri, Cortana, Alexa)

Γενικά η επεξεργασία φυσικής γλώσσας θεωρείται ένα δύσκολο πρόβλημα στην επιστήμη των υπολογιστών λόγω της ίδιας της φύσης της ανθρώπινης γλώσσας. Οι κανόνες που διέπουν τη μετάδοση της πληροφορίες χρησιμοποιώντας φυσική γλώσσα δεν είναι εύκολο να γίνουν κατανοητοί από τον υπολογιστή. Για παράδειγμα, κάποιοι κανόνες μπορεί να είναι εντελώς αφηρημένοι όπως η χρήση σαρκασμού για τη μετάδοση της πληροφορίας. Έτσι, προκειμένου να γίνει κατανοητή η φυσική γλώσσα από τον υπολογιστή θα πρέπει να καταλάβει και τις λέξεις αλλά και πώς συνδέονται οι έννοιες για να παραδοθεί ένα μήνυμα. Η ασάφεια, λοιπόν, της ανθρώπινης γλώσσας είναι αυτή που καθιστά δύσκολη την επεξεργασία της φυσικής γλώσσας και είναι και ένας από τους βασικούς λόγους που ενώ η ιδέα αυτή υπάρχει από πολύ παλιά τα πιο σημαντικά βήματα σε αυτόν τομέα έχουν γίνει τα τελευταία χρόνια.

3.2.2. Ιστορική Αναδρομή

Το 1950 ο Alan Turing έγραψε μια δημοσίευση που περιέγραφε ένα τεστ για μια «σκεπτόμενη» μηχανή. Το τεστ Turing σχεδιάστηκε από τον Alan Turing ως τρόπο να κρίνει την επιτυχία ή αλλιώς μια προσπάθεια παραγωγής ενός υπολογιστή σκέψης. Πιο συγκεκριμένα, βασίστηκε στην ιδέα ότι εάν ένα άτομο που ανακρίνει τον υπολογιστή δε μπορεί να αντιληφθεί εάν πρόκειται για άνθρωπο ή υπολογιστή, τότε όπως είπε ο Turing, πρόκειται για μια έξυπνη μηχανή. Η δοκιμή έχει σχεδιαστεί ως εξής:

Ο ανακριτής έχει πρόσβαση σε δύο άτομα, ένα εκ των οποίων είναι άνθρωπός και το άλλο είναι υπολογιστής. Ο ανακριτής μπορεί να θέσει σε δύο άτομα ερωτήσεις, αλλά δε μπορεί να αλληλεπιδράσει απευθείας μαζί τους. Πιθανώς οι ερωτήσεις εισάγονται σε έναν υπολογιστή μέσω πληκτρολογίου και οι απαντήσεις εμφανίζονται στην οθόνη του υπολογιστή. Ο άνθρωπος σκοπεύει να προσπαθήσει να βοηθήσει τον ανακριτή, αλλά εάν ο υπολογιστής είναι αρκετά έξυπνος, θα πρέπει να είναι σε θέση να ξεγελάσει τον ανακριτή και να είναι αβέβαιο για το ποιος είναι ο υπολογιστής και ποιος είναι ο άνθρωπος. Ο άνθρωπος μπορεί να δώσει απαντήσεις όπως "Είμαι ο άνθρωπος - ο άλλος είναι ο υπολογιστής", αλλά φυσικά, έτσι μπορεί να απαντήσει και ο υπολογιστής. Ο πραγματικός τρόπος με τον οποίο ο άνθρωπος αποδεικνύει την ιδιότητά του είναι να δώσει σύνθετες απαντήσεις που δε θα μπορούσε να κατανοήσει ο υπολογιστής (Benjamin & Gilis, 2021).Λίγο αργότερα από αυτό, το 1952 το μοντέλο των Hodgkin-Huxley έδειξε πώς ο εγκέφαλος χρησιμοποιεί τους νευρώνες για να σχηματίσει ένα ηλεκτρικό δίκτυο. Αυτά τα γεγονότα ενέπνευσαν την ιδέα της τεχνητής νοημοσύνης, του Natural Language Processing και την εξέλιξη των υπολογιστών (Foote, 2019).

Το 1956, ο όρος Τεχνητή Νοημοσύνη χρησιμοποιήθηκε για πρώτη φορά από τη διάσκεψη του John McCarthy στο Dartmouth College, στο Ανόβερο, στο Νιού Χάμσαϊρ .Το 1957, οι Newell και Simon ανακάλυψαν την ιδέα του GPS, του οποίου οι σκοποί, όπως υποδηλώνει το όνομα, ήταν να λύσουν σχεδόν οποιοδήποτε λογικό πρόβλημα. Το πρόγραμμα χρησιμοποίησε μια μεθοδολογία, η οποία βασίζεται στη θεώρηση του προσδιορισμού του τι πρέπει να γίνει και στη συνέχεια να επεξεργαστεί έναν τρόπο να το κάνει. Αυτό λειτουργεί αρκετά καλά για απλά προβλήματα, αλλά οι ερευνητές της τεχνητής νοημοσύνης συνειδητοποίησαν σύντομα ότι αυτό το είδος μεθόδου δε θα μπορούσε να εφαρμοστεί με έναν τόσο γενικό τρόπο. Σε αυτή την εποχή υπήρχε μεγάλη αισιοδοξία για την πρόοδο της Τεχνητής Νοημοσύνης.

Την ίδια χρονιά ο Νόαμ Τσόμσκι ,Αμερικανός καθηγητής στο Τμήμα Γλωσσολογίας και Φιλοσοφίας του Τεχνολογικού Ινστιτούτου της Μασαχουσέτης, δημοσίευσε το βιβλίο του «Syntactic Structures», στο οποίο εξέλιξε προηγούμενες γλωσσικές έννοιες,

καταλήγοντας ότι για να καταλάβει ένας υπολογιστής μια γλώσσα θα έπρεπε να αλλάξει η συντακτική δομή της. Με αυτό σαν στόχο του, ο Τσόμσκι δημιούργησε ένα στυλ γραμματικής που ονομάζεται γραμματική δομή φράσεων, η οποία μεθοδολογικά μετέφρασε τις προτάσεις της φυσικής γλώσσας σε ένα τύπο που μπορεί να χρησιμοποιηθεί από τους υπολογιστές.

Το 1958, ο John McCarthy εισήγαγε τη γλώσσα προγραμματισμού LISP(Locator/Identifier Separation Protocol) που είναι ακόμα σε χρήση σήμερα. Το 1964 σχεδιάστηκε η δακτυλογραφούμενη διαδικασία ELIZA ερωτήσεων και απαντήσεων. Επίσης, την ίδια χρονιά το Αμερικάνικο Εθνικό Συμβούλιο Έρευνας δημιούργησε την Automatic Language Processing Advisory Committee ή ALPAC σε συντομογραφία, μια επιτροπή που επιφορτίστηκε με την αξιολόγηση της προόδου της επεξεργασίας της φυσικής γλώσσας.

Ωστόσο το 1966, το Αμερικάνικο Εθνικό Συμβούλιο Έρευνας και η ALPAC σταμάτησαν για πρώτη φορά τη χρηματοδότηση της έρευνας στην επεξεργασία της φυσικής γλώσσας και στην τεχνητή νοημοσύνη. Κι αυτό, διότι μετά από δώδεκα χρόνια έρευνας και χρηματοδότηση είκοσι εκατομμυρίων δολαρίων οι μηχανικές μεταφράσεις ήταν ακόμα πιο ακριβές από τις μεταφράσεις στο χέρι. Έτσι το 1966, η έρευνα πάνω στην Τεχνητή Νοημοσύνη και στην Επεξεργασία Φυσικής Γλώσσας θεωρήθηκε από πολλούς ότι είχε το οριστικό της τέλος.

Πήρε σχεδόν 14 χρόνια μέχρι το 1980 για την επεξεργασία της φυσικής γλώσσας και την τεχνητή νοημοσύνη να ανακάμψουν. Κατά κάποιο τρόπο, η διακοπή της έρευνας εισήγαγε μια νέα περίοδο φρέσκων ιδεών με την ταυτόχρονη εγκατάλειψη αρχικών ιδεών και εννοιών στη μηχανική μετάφραση. Η ανάμειξη γλωσσολογίας και στατιστικής που ήταν κυρίαρχη μέθοδος στην πρώιμη έρευνα ,αντικαταστάθηκε από καθαρή στατιστική. Τη δεκαετία του 1980, λοιπόν, πραγματοποιήθηκε μια θεμελιώδης ανακατεύθυνση με τις απλές προσεγγίσεις να αντικαθιστούν τη βαθιά ανάλυση και τη διαδικασία αξιολόγησης να γίνεται πιο αυστηρή.

Μέχρι το 1980 η πλειοψηφία των συστημάτων επεξεργασίας φυσικής γλώσσας χρησιμοποιούσε σύνθετους χειρόγραφους κανόνες. Όμως στα τέλη της δεκαετίας του 1980 λόγω της αύξησης της υπολογιστικής δύναμης σε συνδυασμό με τη στροφή προς αλγορίθμους μηχανικής μάθησης, ήρθε μια επανάσταση στον τομέα. Αν και κάποιοι από τους πρώτους αλγορίθμους μηχανικής μάθησης παρήγαγαν αποτελέσματα παρόμοια με τις πιο παλιές μεθοδολογίες που γράφονταν στο χέρι, η έρευνα επικεντρώθηκε στα στατιστικά μοντέλα. Στη δεκαετία του 1990, η δημοτικότητα των στατιστικών μοντέλων για επεξεργασία φυσικής γλώσσας ανέβηκε κατακόρυφα μιας και έγιναν ιδιαίτερα χρήσιμα εκμεταλλευόμενα την μεγάλη ροή διαδικτυακού κειμένου.

Το 2001 , ο Yoshio Bengio και η ομάδα του πρότειναν το πρώτο νευρωνικό γλωσσικό μοντέλο ενώ 10 χρόνια αργότερα το 2011 η εφαρμογή Siri της Apple έγινε γνωστή ως ένας από τους πρώτους βοηθούς τεχνητής νοημοσύνης που χρησιμοποιήθηκε από το γενικό πληθυσμό. Στην ουσία η Siri, με την αυτοματοποιημένη αναγνώριση φωνής μεταφράζει τις λέξεις σε ψηφιακές έννοιες. Στη συνέχεια το σύστημα φωνητικών εντολών ταιριάζει αυτές τις έννοιες σε προορισμένες εντολές που ενεργοποιούν κάποιες δράσεις.

Χρησιμοποιώντας, όμως, τεχνικές μηχανικής μάθησης, δεν χρειάζεται οι ανθρώπινες λέξεις να ταιριάζουν ακριβώς σε προορισμένες εκφράσεις. Οι ήχοι θα πρέπει απλά να είναι αρκετά κοντά και το σύστημα θα μεταφράσει το νόημα σωστά. Έτσι, οι μηχανές επεξεργασίες φυσικής γλώσσας μπορούν να βελτιώσουν σημαντικά την ακρίβεια στις μεταφράσεις τους και να αυξήσουν το λεξιλόγιο του συστήματος, δηλαδή τις λέξεις ή εκφράσεις που το σύστημα αναγνωρίζει.

Σήμερα, νευρωνικά μοντέλα θεωρούνται τεχνολογία αιχμής στην έρευνα και ανάπτυξη στους κλάδους της επεξεργασίας φυσικής γλώσσας που ασχολούνται με την αναγνώριση κειμένου και παραγωγή λόγου. Θεωρείται πιθανό να αναπτυχθούν συστήματα ικανά να πιάνουν συζήτηση και να ακούγονται σαν άνθρωποι κάνοντας ερωτήσεις, υποθέσεις και δίνοντας απαντήσεις. Ωστόσο, η σημερινή μας πρόοδος στην τεχνητή νοημοσύνη δεν έχει φτάσει ακόμα σε αυτό το επίπεδο (Foote, 2019).

3.3. Νέοι Μηχανισμοί Διάδοσης Fake News και Αντίμετρα

3.3.1. Τεχνολογία deepfake

Βεβαίως με όλη αυτήν την εξέλιξη της τεχνολογίας φαίνεται να δημιουργείται πρόσφορο έδαφος σε κακοπροαίρετους χρήστες του διαδικτύου να αναπτύξουν και να εξελίξουν ακόμα περισσότερους τρόπους διάδοσης Fake News. Το 2017 η Samantha Cole, δημοσιογράφος του Motherboard ανακάλυψε μια νέα τάση στο διαδίκτυο, τη λεγόμενη τεχνολογία Deepfake (Dickson, 2019).Ένας χρήστης του Reddit με όνομα «deepfakes» δημοσιοποιούσε ψεύτικο πορνογραφικό υλικό χρησιμοποιώντας έναν αλγόριθμο τεχνητής νοημοσύνης ο οποίος άλλαζε τα πρόσωπα αληθινών πρωταγωνιστών σε πορνογραφικό υλικό με αυτά διάσημων προσωπικοτήτων. Η δημοσιογράφος προειδοποίησε για την επικινδυνότητα του συγκεκριμένου φαινομένου, αλλά ένα χρόνο αργότερα ήδη η κατάσταση είχε αποκτήσει τρομακτικές διαστάσεις, μιας και εμφανίστηκαν μια σειρά από εύκολα προσβάσιμες εφαρμογές που μπορούσαν να «αφαιρέσουν» τα ρούχα σε οποιαδήποτε φωτογραφία γυναίκας. Αυτό οδήγησε και το Ντάνιελ Σίτρον, καθηγητή Νομικής στο πανεπιστήμιο της Βοστώνης να πει ότι η τεχνολογία Deepfake εργαλειοποιείται εναντίον της γυναίκας.

Όπως λέει ο Χένρι Άιντερ, επικεφαλής της έρευνας της εταιρείας ανίχνευσης Deeptrace στο Amsterdam, ακόμα και μέχρι σήμερα τα deepfakes συντριπτική τους πλειοψηφία χρησιμοποιούνται για ψεύτικο πορνογραφικό υλικό (Sample, 2020).Ωστόσο, υπάρχει ο κίνδυνος ότι πολιτικά deepfakes μπορούν να αρχίσουν να γίνονται κυρίαρχα και να δημιουργήσουν χάος στην πολιτική ζωή. Για παράδειγμα αν κάποιος θέλει σήμερα να ακούγεται η φωνή του σαν τη φωνή κάποιου γνωστού πολιτικού όπως του Ντόναλντ Τραμπ ή του Μπαράκ Ομπάμα μπορεί πολύ εύκολα να το κάνει. Το μόνο που έχει να κάνει είναι να χρησιμοποιήσει μια εφαρμογή για να ηχογραφήσει μια πρόταση και στη

συνέχεια να ακούσει τι είπε στη φωνή κάποιου διάσημου προσώπου. Καταλαβαίνουμε ότι αυτό παλιότερα θα μπορούσε να επιτευχθεί μόνο αν κάποιος μπορούσε να μιμηθεί φυσικά τη φωνή κάποιου διάσημου προσώπου. Σήμερα, όμως, η ευκολία με την οποία ένας οποιοσδήποτε χρήστης του διαδικτύου μπορεί να κάνει κάτι τέτοιο δημιουργεί προβλήματα σίγουρα ηθικά αλλά και ζητήματα απειλής της δημοκρατίας. Αρκεί να φανταστούμε την μορφή σύγχρονων εκλογών χωρίς να υπάρχει αξιόπιστη πηγή πληροφοριών, όπως οπτικά ή ακουστικά αποσπάσματα.

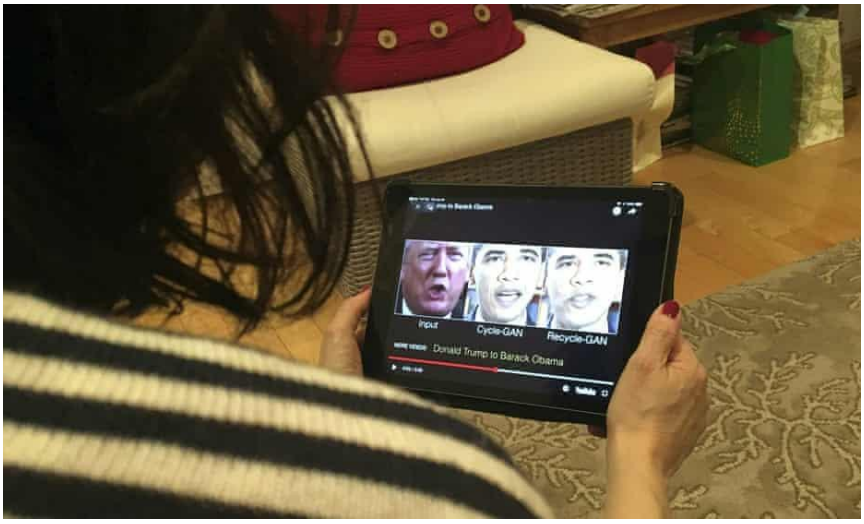
Εικόνα 3: Σύγκριση ενός αυθεντικού και ενός deepfake βίντεο του Ρώσσου Προέδρου Βλαντιμίρ Πούτιν. Φωτογραφία: Alexandra Robinson/AFP Πηγή: https://www.theguardian.com/technology/2020/jan/13/what-are-deepfakes-and-how-can-you-spot-them

Σύμφωνα με τους Karen Hao και Douglas Heaven στο άρθρο τους «The year deepfakes went mainstream» (Hao & Heaven, 2020), το Φεβρουάριο του 2020, κάτι τέτοιο μάλιστα προσπάθησε ο Ινδός πολιτικός Μανόι Τιβάρι χρησιμοποιώντας deepfakes σε ένα προεκλογικό βίντεο προκειμένου να φανεί ότι μιλάει μια συγκεκριμένη ινδική διαλεκτό που μιλούσαν οι ψηφοφόροι στους οποίους στόχευε. Υπάρχουν, βέβαια, κατά τους Karen Hao και Douglas Heaven και άλλες περιπτώσεις deepfakes στην πολιτική που δεν δημιουργήθηκαν με κακοπροαίρετο σκοπό όπως δυο διαφημίσεις που φτιάχτηκαν το Σεπτέμβριο του 2020 πριν τις αμερικάνικες εκλογές με ψεύτικες εκδοχές του Ρώσου προέδρου Βλαντιμιρ Πούτιν και του αρχηγού της Βόρειας Κορέας Κιμ Γιονγκ Ουν να λένε ότι κανείς τους δεν χρειάζεται να παρέμβει στις Αμερικάνικες Εκλογές, γιατί η Αμερική θα καταστρέψει τη δημοκρατία της από μόνη της. Αυτά τα βίντεο ήταν μέρος μιας εκστρατείας της μη κομματικής ομάδας RepresentUs με σκοπό να ταρακουνήσει τους Αμερικάνους για την ευθραυστότητα της δημοκρατίας σε μια περίοδο που ο Πρόεδρος Τραμπ άφηνε υπόνοιες περί μη ομαλής μετάβασης του νέου Προέδρου στην εξουσία σε περίπτωση που χάσει τις εκλογές.

3.3.2. Deepfakes και λειτουργίες

Το Deepfake υλικό στην ουσία πρόκειται για βίντεο, φωτογραφίες και ηχητικά που έχουν παραχθεί από αλγορίθμους τεχνητής νοημοσύνης. Κύριο ρόλο στην διασπορά των ψευδών ειδήσεων μέσω deepfake και πιο διαδεδομένο αποτελούν κυρίως τα βίντεο και οι φωτογραφίες. Θα αναλύσουμε τη βασική διαδικασία που συνήθως ακολουθείται μιας και η σε βάθος ανάλυση των συγκεκριμένων τεχνολογιών δεν αποτελούν σκοπό της παρούσας εργασίας.

Όπως περιγράφει ο Ian Sample στο άρθρο του «What are deepfakes – and how can you spot them?» (Sample, 2020) υπάρχουν μόνο μερικά βήματα που χρειάζεται κανείς να κάνει για να δημιουργήσει ένα βίντεο με εναλλαγές προσώπων. Αρχικά τροφοδοτείται ένας αλγόριθμος τεχνητής νοημοσύνης ο οποίος ονομάζεται encoder με χιλιάδες φωτογραφίες των δύο προσώπων. Ο encoder βρίσκει και μαθαίνει τις ομοιότητες μεταξύ των δύο προσώπων και τις μειώνει στα κοινά τους χαρακτηριστικά συμπιέζοντας τις φωτογραφίες κατά τη διαδικασία. Ένας δεύτερος αλγόριθμος τεχνητής νοημοσύνης που ονομάζεται decoder μαθαίνει να επαναφέρει τα πρόσωπα από τις συμπιεσμένες φωτογραφίες. Επειδή τα πρόσωπα είναι διαφορετικά, εκπαιδεύεται ένας decoder να επαναφέρει το πρόσωπο Α και ένας decoder να επαναφέρει το πρόσωπο Β. Για να πραγματοποιηθεί η ανταλλαγή προσώπων, απλά εισάγονται οι φωτογραφίες του encoder στο «λάθος» decoder. Για παράδειγμα, μια συμπιεσμένη φωτογραφία του προσώπου Α εισάγεται στο decoder που εκπαιδεύτηκε με το πρόσωπο Β. Τότε ο decoder ανακατασκευάζει το πρόσωπο Β με τις εκφράσεις και τον προσανατολισμό του προσώπου Α.



Εικόνα 4: Μια γυναίκα βλέπει ένα deepfake video του Ντόναλντ Τραμπ και του Μπαράκ Ομπάμα.Φωτογραφία:Rob Lever/AFP μέσω Getty Images.Πηγή: https://www.theguardian.com/technology/2020/jan/13/what-are-deepfakes-and-how-can-you-spot-them

Ένας άλλος τρόπος να δημιουργηθούν πλαστές φωτογραφίες προσώπων είναι αυτό που ονομάζεται Generative Adversarial Network (GAN). Σε αυτή τη διαδικασία χρησιμοποιούνται δύο αλγόριθμοι τεχνητής νοημοσύνης . Ο πρώτος αλγόριθμος, γνωστός και ως «generator» τροφοδοτείται με τυχαίο θόρυβο πληροφορίας και τον μετατρέπει σε μια φωτογραφία. Αυτή η συνθετική φωτογραφία στη συνέχεια μαζί με άλλες αληθινές φωτογραφίες, διασημοτήτων για παράδειγμα, εισάγονται σε έναν δεύτερο αλγόριθμο, γνωστός και ως «discriminator». Αρχικά, οι συνθετικές φωτογραφίες δεν μοιάζουν καθόλου με πρόσωπα. Αλλά επαναλαμβάνοντας αυτή τη διαδικασία πάρα πολλές φορές και με ανατροφοδότηση κατά την εκτέλεση, ο generator και ο discriminator βελτιώνονται. Κάποια στιγμή ο generator θα αρχίσει να παράγει εντελώς ρεαλιστικά πρόσωπα ανθρώπων που δεν υπάρχουν.

3.3.3. Αντιμετώπιση Deepfake

Τα deepfake βίντεο είναι αρκετά δύσκολο να ανιχνευτούν με το μάτι γιατί μοιάζουν αρκετά ρεαλιστικά. Ο Siwei Lyu, καθηγητής Επιστήμης Υπολογιστών στο University

at Albany-State University of New York, έχει ερευνήσει τα τελευταία χρόνια το ζήτημα της ανίχνευσης των deepfake βίντεο και έχει καταφέρει να δώσει κάποιες λύσεις.

Το 2018 με τη δημοσίευση του «ExposingAI Created Fake Videos by Detecting Eye Blinking» (Li, Chang, & Lyu, 2018) βρέθηκε μια πρώτη απάντηση στα βίντεο αυτά. Όπως χαρακτηριστικά λέει στο άρθρο του «Detecting deepfake videos in the blink of an eye» (Lyu S. , 2018) οι υγιείς ενήλικες άνθρωποι ανοιγοκλείνουν τα μάτια τους κάθε περίπου 2 έως 10 δευτερόλεπτα. Όμως επειδή στα deepfake βίντεο οι αλγόριθμοι που τα κατασκευάζουν δεν έχουν τροφοδοτηθεί με αρκετές φωτογραφίες προσώπων με κλειστά ματιά παρατήρησε ότι τα μάτια των προσώπων ανοιγοκλείνουν με διαφορετική συχνότητα. Έτσι στην έρευνα του χρησιμοποίησε τεχνικές μηχανικής μάθησης για να εξετάσει ακριβώς αυτή τη συχνότητα και αν είναι αποκλίνουσα από τη φυσιολογική το βίντεο να θεωρείται πλαστό.

Η τεχνολογία, όμως, εξελίσσεται πολύ γρήγορα και μαζί της ο ανταγωνισμός ανάμεσα στην παραγωγή και στην ανίχνευση πλαστών βίντεο. Έτσι, οι παραγωγοί τέτοιων βίντεο προσπέρασαν αυτό το εμπόδιο τροφοδοτώντας τους αλγορίθμους τους και με φωτογραφίες προσώπων με κλειστά μάτια ή ακόμα καλύτερα με ολόκληρα βίντεο και κατάφεραν να κάνουν τα βίντεο τους να μοιάζουν ακόμα πιο ρεαλιστικά.

Αυτό ενέπνευσε τη συνέχεια της έρευνας στην ανίχνευση πλαστών βίντεο. Ο Siwei Lyu και η ομάδα του με δύο ακόμα δημοσιεύσεις «Exposing DeepFake Videos By Detecting Face Warping Artifacts» (Li & Lyu, 2018), «Exposing Deep Fakes Using Inconsistent Head Poses» (Yang, Li, & Lyu, 2018) περιέγραψαν τρόπους να ανιχνευτούν τα πλαστά βίντεο από κάποια ελαττώματα τους που δεν είναι εύκολο να διορθωθούν από τους παραγωγούς τους.

Συγκεκριμένα, όπως εξηγεί ο Siwei Lyu στο άρθρο του «Detecting deepfakes by looking closely reveals a way to protect against them» (Lyu S. , 2019), οι deepfake αλγόριθμοι δεν είναι ακόμα ικανοί να κατασκευάσουν πρόσωπα στις τρεις διαστάσεις. Στην ουσία, δημιουργούν μια φωτογραφία δυο διαστάσεων του προσώπου και μετά προσπαθούν να την περιστρέψουν και να της αλλάξουν το μέγεθος για να ταιριάζει στην κατεύθυνση που φαίνεται ότι το πρόσωπο κοιτάει. Έτσι, η ερευνητική ομάδα του Lyu σχεδίασε έναν αλγόριθμο που υπολογίζει τις κατευθύνσεις της μύτης και του κεφαλιού του προσώπου. Στα αληθινά βίντεο αυτές οι κατευθύνσεις σχεδόν ευθυγραμμίζονται ενώ στα πλαστά βίντεο πολύ συχνά αυτό δεν συμβαίνει.

Εικόνα 5: Όταν ένας υπολογιστής βάζει το πρόσωπο του Νίκολας Κέιτζ στο κεφάλι του Έλον Μασκ , μπορεί να μην ευθυγραμμίσει το κεφάλι και το πρόσωπο σωστά.Πηγή: https://theconversation.com/detecting-deepfakes-by-looking-closely-reveals-a-way-to-protect-against

Αυτή τη στιγμή η ίδια ερευνητική ομάδα δουλεύει σε μια νέα ιδέα που ίσως θα μπορέσει να εφαρμοστεί στο μέλλον για να επιτευχθεί μια γενικότερη προστασία. Η ουσία της ιδέας αυτής είναι οποιαδήποτε φωτογραφία ανεβαίνει στο διαδίκτυο να ερωτάται ο χρήστης αν θέλει να προστατευτεί η φωτογραφία του από αναγνώριση προσώπου. Στην περίπτωση της θετικής απάντησης η φωτογραφία να υπόκειται σε κάποιου είδους επεξεργασία προσθέτοντας της «θόρυβο», έτσι ώστε όταν ένας αλγόριθμος που προσπαθεί να αναγνωρίσει το πρόσωπο σε μια φωτογραφία να μην μπορεί να το κάνει. Έτσι θα γίνεται αρκετά δύσκολο για έναν αλγόριθμο να συλλέγει σε μια βιβλιοθήκη πρόσωπα που αργότερα θα μπορεί να χρησιμοποιήσει για την εκπαίδευση των αλγορίθμων παραγωγής πλαστών deepfake βίντεο. Το μόνο σίγουρο είναι ότι αυτή τη στιγμή οι πλατφόρμες έχουν κάποια σημαντικά εργαλεία που μπορούν να αξιοποιήσουν έτσι ώστε να καταπολεμήσουν τέτοιου είδους φαινόμενα.

3.3.4. Αυτοματοποιημένη παραγωγή κειμένου

Όπως αναλύθηκε νωρίτερα στο κεφάλαιο η εξέλιξη στον κλάδο της επεξεργασίας φυσικής γλώσσας και στην παραγωγή κειμένου είναι ραγδαία. Αυτή η εξέλιξη μπορεί να ανοίγει το δρόμο για περαιτέρω αναβάθμιση ήδη υπαρχόντων εφαρμογών όπως μεταφράσεις ή αυτόματη εξυπηρέτηση πελατών μέσω κειμένου ή και ανάπτυξη ακόμα πιο νέων ιδεών και εφαρμογών, όμως είναι πραγματικά πολύ επικίνδυνος ο τρόπος με τον οποίο μπορεί να χρησιμοποιηθεί αυτή η τεχνολογία.

Αρκεί να αναφερθεί ότι οι δημιουργοί του μοντέλου GPT2 το Φεβρουάριο του 2019, του πρώτου συστήματος τεχνητής νοημοσύνης που παράγει κείμενο, πήραν την ασυνήθιστη απόφαση να μην δημοσιεύσουν αρχικά την έρευνα τους υπό το φόβο της κακής χρήσης του εργαλείου αυτού (Hern, 2019). Μάλιστα δημιούργησαν και δημοσιοποίησαν ένα ανοιχτού κώδικα μοντέλο ή σε πιο εύκολη μορφή το GPTrue or False που λειτουργεί ως επέκταση στο Google Chrome, το οποίο μπορεί να ανιχνεύει κείμενα που έχουν παραχθεί από μηχανή και συγκεκριμένα από το ίδιο το GPT2 (Ambalina, 2020). Λίγο καιρό αργότερα δημοσιοποίησαν και το GPT2 με σκοπό να

δώσουν το έναυσμα σε περισσότερη έρευνα στην ανίχνευση κειμένων που έχουν παραχθεί από συστήματα τεχνητής νοημοσύνης.

Από τότε έχουν δημιουργηθεί διάφορα άλλα μοντέλα παραγωγής κειμένου με καλύτερο μέχρι στιγμής σε ακρίβεια το GROVER (Fagni, Falchi, Gambini, Martella, & Tesconi, 2021). Βέβαια στη δημοσίευση του GROVER (Zellers, και συν., 2019) προτάθηκε η άποψη ότι καλύτερο ανιχνευτικό σύστημα για ένα κείμενο που παράχθηκε από μηχανή είναι το ίδιο το σύστημα που το δημιούργησε. Ωστόσο, αυτή η άποψη αποδείχτηκε λανθασμένη μιας και η OpenAI, η ομάδα που δημιούργησε το GPT2, κατάφερε να πετύχει μεγαλύτερη ακρίβεια σε κείμενα που είχαν παραχθεί από το GPT2, χρησιμοποιώντας διαφορετικό ανιχνευτικό σύστημα (Fagni, Falchi, Gambini, Martella, & Tesconi, 2021).

Το σίγουρο είναι από εδώ και στο εξής ότι ανοίγεται μεγάλο πεδίο έρευνας στο συγκεκριμένο τομέα, ειδικά αν αναλογιστεί κανείς ότι μόλις φέτος το 2021 δημιουργήθηκε το πρώτο σύνολο δεδομένων από fake news που πράγματι δημοσιεύτηκαν στο twitter, κάποια γραμμένα από ανθρώπους και κάποια άλλα από μηχανές (Fagni, Falchi, Gambini, Martella, & Tesconi, 2021). Βέβαια, όλα αυτά βρίσκονται σε αρκετά πρώιμο στάδιο έρευνας και τα περισσότερα fake news με τα οποία ερχόμαστε αντιμέτωποι είναι γραμμένα από ανθρώπους (Zellers, και συν., 2019), (Fagni, Falchi, Gambini, Martella, & Tesconi, 2021) και για αυτό στη συνέχεια της εργασίας μας θα επικεντρωθούμε σε αυτά εμβαθύνοντας στα μαθηματικά ενός αρκετά αποδοτικού στατιστικού μοντέλου, όπως αυτό της λογιστικής παλινδρόμησης.

ΚΕΦΑΛΑΙΟ 4
ΣΤΑΤΙΣΤΙΚΗ ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΚΑΙ ΜΟΝΤΕΛΟ ΛΟΓΙΣΤΙΚΗΣ ΠΑΛΙΝΔΡΟΜΗΣΗΣ

4.1. Εισαγωγή

Η εκτίμηση και η περιγραφή της σχέσης εξάρτησης μεταξύ μεταβλητών αποτελεί βασικό στόχο σε πολλές επιστήμες. Μάλιστα, αυτή η σχέση σε πολλές φορές δεν είναι συναρτησιακή, αλλά στοχαστική. Τέτοιου είδους φαινόμενα είναι για παράδειγμα η ανταπόκριση ενός ασθενή σε μια συγκεκριμένη δόση φαρμάκου ή η κατανάλωση βενζίνης σε ένα συγκεκριμένο ταξίδι. Η σχέση, λοιπόν, ανάμεσα σε αυτές τις μεταβλητές είναι ανάγκη να εκφραστούν από ένα στατιστικό μοντέλο , το οποίο να ενσωματώνει με κάποιο τρόπο την έννοια της αβεβαιότητας και ,πιο συγκεκριμένα, την έννοια των «τυχαίων σφαλμάτων».

Ο κλάδος της Στατιστικής που εξετάζει τη σχέση μεταξύ δύο ή περισσότερων μεταβλητών με απώτερο σκοπό την πρόβλεψη μια από αυτές μέσω των άλλων ονομάζεται Ανάλυση Παλινδρόμησης ή με τον αγγλικό όρο Regression Analysis (Κούτρας, 2011) ενώ η σχέση της εξαρτημένης και των ανεξάρτητων μεταβλητών ονομάζεται εξίσωση παλινδρόμησης ή μοντέλο παλινδρόμησης.

Η μαθηματική έκφραση ενός τέτοιου μοντέλου δύο μεταβλητών αν θεωρήσουμε Y την εξαρτημένη μεταβλητή και X την ανεξάρτητη είναι :

$$Y=f(X)+\varepsilon, \text{ όπου } \varepsilon \sim (0,\sigma^2)$$

4.2. Απλό Γραμμικό Μοντέλο Παλινδρόμησης

Για να καταλάβουμε, όμως, καλύτερα τον τρόπο που δουλεύει ένα μοντέλο παλινδρόμησης ας αναλύσουμε πρώτα το απλό γραμμικό μοντέλο. Το μοντέλο αυτό περιλαμβάνει δύο μεταβλητές την ανεξάρτητη X και την εξαρτημένη Y , οι οποίες συνδέονται μεταξύ τους με τη γραμμική συνάρτηση παλινδρόμησης (Καρώνη & Οικονόμου, 2010).Σκοπός είναι η προσαρμογή μιας ευθείας γραμμής, η οποία επεξηγεί όσο το δυνατόν καλύτερα η συμπεριφορά των δεδομένων. Μια τέτοια ευθεία θα έχει τη μορφή :

$$E(Y \mid X = x) = b_0 + b_1 x \qquad (4.1)$$

Στην προαναφερθείσα σχέση, E(Y|X=x) η αναμενόμενη τιμή της μεταβλητής Y για συγκεκριμένη τιμή x της μεταβλητής X και b_0,b_1 οι συντελεστές παλινδρόμησης. Η προσαρμογή της καλύτερης ευθείας, δηλαδή η καλύτερη δυνατή εκτίμηση των παραμέτρων, γίνεται λαμβάνοντας υπόψη τις n ανεξάρτητες παρατηρήσεις(x_i,y_i), i=1,...,n, που έχουμε στη διάθεση μας προς επεξεργασία.

Τα σημεία (x_i,y_i),i=1…,n, είναι πιθανό να διαφέρουν από τα σημεία (x_i, $\hat{y}_i$), όπου

$$\hat{y}_i = \hat{b}_0 + \hat{b}_1 x_i$$

είναι η εκτίμηση της τιμής της τυχαίας μεταβλητής Y με βάση το απλό γραμμικό μοντέλο που προσαρμόστηκε στα δεδομένα του προβλήματος και $\hat{b}_0, \hat{b}_1$ οι εκτιμήσεις των παραμέτρων του μοντέλου.

Οι παρατηρήσεις y_i δίνονται από τη σχέση :

$$y_i=E(Y_i|X_i=x_i)+\varepsilon_i=b_0+b_1x_i+\varepsilon_i \quad (4.2)$$

Το ε_i ονομάζεται τυχαίο σφάλμα και παριστάνει για δοθείσα τιμή x_i την κατακόρυφη απόκλιση της τιμής y_i από την ευθεία της συνάρτησης παλινδρόμησης. Αντίστοιχα, η διαφορά :

$$e_i=y_i - \hat{y}_i = y_i - \hat{b}_0 - \hat{b}_1 x_i$$

αποτελεί την κατακόρυφη απόκλιση του y_i από την ευθεία της εκτιμημένης συνάρτησης παλινδρόμησης και ονομάζεται υπόλοιπο (residual). Τα e_i μπορούν να θεωρηθούν ως οι εκτιμήσεις των άγνωστων τυχαίων σφαλμάτων ε_i.

Η προσαρμογή του μοντέλου, δηλαδή η εκτίμηση των παραμέτρων b_0 και b_1, γίνεται με τη γνωστή μέθοδο των ελαχίστων τετραγώνων που για λόγους συντομίας δεν θα αναλυθεί.

Τέλος, από τη γνώση πλέον της ευθείας παλινδρόμησης μπορεί να γίνεται πρόβλεψη της μεταβλητής Y για οποιαδήποτε τιμή της μεταβλητής X.

4.3. Λογιστικό Μοντέλο Παλινδρόμησης

Το μοντέλο της Λογιστικής Παλινδρόμησης (Logistic Regression) ανήκει στην κατηγορία των γενικευμένων γραμμικών μοντέλων. Η βασική διαφορά μεταξύ λογιστικής και γραμμικής παλινδρόμησης βασίζεται στο είδος της εξαρτημένης μεταβλητής, η οποία στην πρώτη μπορεί να είναι κατηγορική (ονομαστική ή τακτική) ενώ στη δεύτερη αποκλειστικά ποσοτική. Επίσης, σε αντίθεση με την κλασική γραμμική παλινδρόμηση όπου η εκτίμηση των παραμέτρων b_o και b_1 γίνεται με τη μέθοδο των ελαχίστων τετραγώνων, κατά τη λογιστική παλινδρόμηση η εκτίμηση των παραμέτρων γίνεται με τη μέθοδο εκτίμησης μέγιστης πιθανοφάνειας, δηλαδή επιλέγονται οι πιο πιθανοφανείς τιμές των παραμέτρων, προκειμένου να οδηγήσουν στα παρατηρούμενα αποτελέσματα.

Διακρίνονται τρεις τύποι λογιστικής παλινδρόμησης ανάλογα με την ιδιαίτερη φύση της εξαρτημένης κατηγορικής μεταβλητής (repository Kallipos):

1. Δίτιμη ή δυαδική (Binary) ή διμερής εξαρτημένη μεταβλητή. Αποτελείται από δυο κατηγορίες, όπως παραδείγματος χάρη οι εκβάσεις αποτυχία/αποτυχία, Ναι/Όχι.
2. Τακτική (ordinal) μεταβλητή. Η εξαρτημένη μεταβλητή αποτελείται από τρεις ή περισσότερες κατηγορίες μεταξύ των οποίων ισχύει η έννοια της ανισότητας, όπως π.χ σε μια ερώτηση της κλίμακας είμαι πολύ ευχαριστημένος, αρκετά ευχαριστημένος, λίγο ευχαριστημένος, καθόλου ευχαριστημένος.

3. Ονομαστική (Nominal) ή πολυωνυμική (polynomial) ή πολυχοτομική (polychotomus) ή κατηγορική αδιαβάθμητη (non-ordered categorical) ή πολυμερής μεταβλητή απόκρισης. Περιέχει τρεις ή περισσότερες κατηγορίες χωρίς κάποια φυσική διαβάθμιση, όπως π.χ. ο χαρακτηρισμός ενός του χρώματος αντικειμένων ως ερυθρού, πράσινου, κίτρινου κλπ.

Η λογιστική παλινδρόμηση έχει ευρεία εφαρμογή σε πολλά επιστημονικά πεδία και σε διάφορες εφαρμογές. Χαρακτηριστικά χρησιμοποιείται στην πρόβλεψη της :

- Εμφάνισης ή μη μιας νόσου
- Επιλογής ενός πολιτικού κόμματος
- Πιθανότητας αποτυχίας μιας διεργασίας παραγωγής προϊόντος σε ένα εργοστάσιο τροφίμων
- Πιθανότητας αθέτησης από δανειολήπτη αποπληρωμής του δανείου του

Εκτός των προαναφερθέντων περιπτώσεων, μπορεί να βρει χρήση σε πολλαπλές συνθήκες με διαφορετικούς σκοπούς. Η θεμελίωση του μοντέλου αυτού ακολουθεί στο επόμενο κεφάλαιο.

4.4. Μαθηματική θεμελίωση του μοντέλου

Είναι σημαντικό να κατανοήσουμε ότι για ένα πρόβλημα δυαδικής εξαρτημένης μεταβλητής η αριθμητική τιμή της μεταβλητής Y είναι αυθαίρετη και ως εκ τούτου δεν παρουσιάζει κάποιο ιδιαίτερο ενδιαφέρον. Αυτό που είναι πραγματικά ενδιαφέρον είναι εάν η ταξινόμηση των περιπτώσεων σε μια από τις 2 κατηγορίες της εξαρτημένης μεταβλητής μπορεί να προβλεφθεί από τις ανεξάρτητες μεταβλητές. Αντί, δηλαδή να προσπαθούμε να προβλέψουμε μια αυθαίρετη τιμή , είναι πιο χρήσιμο να προσπαθούμε να προβλέψουμε την πιθανότητα η περίπτωση μας να ταξινομείται στη μια ή στην άλλη κατηγορία (Menard, 2010).Ως εκ τούτου, συμβολίζουμε την πιθανότητα ένα άρθρο να ταξινομείται στην πρώτη κατηγορία (real news) με P(Y=0) ενώ την πιθανότητα να ταξινομείται στη δεύτερη κατηγορία (fake news) με P (Y=1). Η P(Y=1) ονομάζεται και πιθανότητα «επιτυχίας» και στο δικό μας πρόβλημα την ταυτίζουμε με την πιθανότητα ένα άρθρο να είναι fake news μιας και αυτό αφορά η αναζήτηση μας.

4.4.1. Μετασχηματισμός συμπληρωματικών πιθανοτήτων

Θα μπορούσαμε να προσπαθήσουμε να μοντελοποιήσουμε την πιθανότητα Y=1 σαν $P(Y=1)=b_0+b_1X$, αλλά θα είχαμε το πρόβλημα ότι οι προβλεπόμενες τιμές μπορεί να έπαιρναν τιμές μικρότερες του 0 ή μεγαλύτερες του 1.Ένα βήμα μπροστά είναι η αντικατάσταση της πιθανότητας Y=1 με το λόγο συμπληρωματικών πιθανοτήτων (odds) Y=1. Τα odds Y=1 ή πιο σύντομα Ω(Y=1) είναι ο λόγος της πιθανότητας «επιτυχίας» προς την πιθανότητα «αποτυχίας». Με μαθηματικά:

$$\Omega(Y=1)=\frac{P(Y=1)}{1-P(Y=1)}$$

Τα odds δεν έχουν πάνω όριο άλλα θα πρέπει να έχουν κάτω όριο το 0 . Όμως αυτός ο περιορισμός δεν ικανοποιείται γιατί με το μετασχηματισμό $\Omega = b_0+b_1X$ παίρνουν οποιαδήποτε τιμή. Έτσι χρειαζόμαστε ένα νέο μετασχηματισμό.

4.4.2. Μετασχηματισμός LOGIT

Λογαριθμίζοντας το λόγο των συμπληρωματικών πιθανοτήτων έχουμε :

$$\ln\left(\frac{P(\mathrm{Y}=1)}{1-P(Y=1)}\right)$$

Αυτή η ποσότητα ονομάζεται logit του Y και προσεγγίζει το $-\infty$ όταν τα odds τείνουν στο 0 ενώ προσεγγίζει το $+\infty$ όταν τα odds τείνουν στο $+\infty$.

Αν λοιπόν χρησιμοποιήσουμε το φυσικό λογάριθμο των odds σαν εξαρτημένη μεταβλητή όπως θα δούμε παρακάτω , δεν αντιμετωπίζουμε πλέον το πρόβλημα των δυνατών τιμών για την πιθανότητα. Έτσι η εξίσωση ανάμεσα σε εξαρτημένη και ανεξάρτητες μεταβλητές γίνεται :

$$logit(Y)=b_0+b_1X_1+b_2X_2+\ldots b_\kappa X_\kappa \qquad (4.3)$$

Μπορούμε να μετατρέψουμε ξανά το logit(Y) σε odds λαμβάνοντας υπόψη ότι $\Omega(Y=1)=e^{logit(Y)}$.Άρα:

$$\Omega(Y=1)=e^{\ln\{\Omega(Y=1)\}}=e^{b_0+b_1X_1+b_2X_2+b_\kappa X_\kappa}$$

4.4.3. Σιγμοειδής λογιστική συνάρτηση

Μετατρέπουμε τώρα τα odds σε πιθανότητες από τη σχέση (4.6) και επιλύουμε ως προς P(Y=1) καταλήγοντας στην

$$\mathrm{P}(\mathrm{Y}=1)=\frac{\Omega(\mathrm{Y}=1)}{1+\Omega(\mathrm{Y}=1)}$$

Έτσι προκύπτει η εξίσωση:

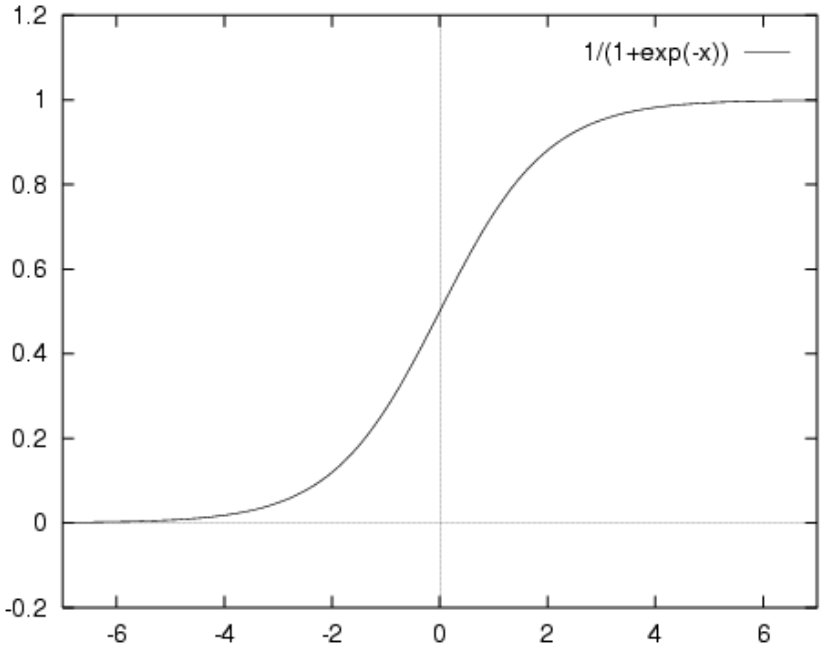
$$P(Y=1)=\frac{e^{b_0+b_1X_1+b_2X_2\ldots+b_kX_k}}{1+e^{b_0+b_1X_1+b_2X_2\ldots+b_kX_k}}$$

Αν θέσουμε $z= e^{\alpha+\beta_1X_1+\beta_2X_2+\beta_3X_3+\ldots\beta_\kappa X_\kappa}$ τότε η συνάρτηση που μας δίνει την πιθανότητα «επιτυχίας» είναι η :

$$\sigma(z)=\frac{e^z}{e^z+1}=\frac{1}{1+e^{-z}} \qquad (4.4)$$

η οποία λέγεται και σιγμοειδής συνάρτηση επειδή η καμπύλη της σχηματίζει το γράμμα S.

Βλέπουμε πράγματι ότι το πεδίο τιμών της σ(z) είναι το (0,1) και άρα ικανοποιεί τους περιορισμούς της πιθανότητας. Συγκεκριμένα, αν $z \to -\infty$ τότε $\sigma(z) \to 0$ ενώ αν $z \to +\infty$ τότε $\sigma(z) \to 1$.



Γράφημα 1:Γραφική παράσταση της σιγμοειδούς συνάρτησης Πηγή: https://computing.dcu.ie/~humphrys/Notes/Neural/sigmoid.html

4.5. Μέθοδος Εκτίμησης Μέγιστης Πιθανοφάνειας

Όπως αναφέρθηκε και παραπάνω, η μεταβλητή Y έχει μόνο δύο δυνατές τιμές,0 και 1.Έτσι η Y είναι τυχαία μεταβλητή της κατανομής Bernoulli , δηλαδή Y~B(p) ,όπου p=P(Y=1). Η συνάρτηση πιθανότητας για μεταβλητή y της κατανομής Bernoulli είναι:

$$f(y)=P(Y=y)=p^y(1-p)^{1-y}, \text{ όπου } y=0,1$$

Η μέση τιμή μ της Y είναι :

$$\mu=E(Y)=0 \cdot f(0)+1 \cdot f(1)=p$$

Και η διακύμανση της:

$$\sigma^2=E(Y^2)-[E(Y)]^2=0^2 \cdot f(0)+1^2 \cdot f(1) -p^2=p-p^2=p \cdot (1-p)$$

Οι συντελεστές $\beta_1,\beta_2,\beta_3 \ldots \beta_\kappa$ της συνάρτησης πιθανότητας «επιτυχίας» υπολογίζονται από τη μέθοδο εκτίμησης μέγιστης πιθανοφάνειας (ΕΜΠ).Η μέθοδος αυτή επιλέγει τους συντελεστές β έτσι ώστε να μεγιστοποιείται η συνάρτηση πιθανοφάνειας (Likelihood function) L (Καρώνη & Οικονόμου, 2010).

Η συνάρτηση πιθανοφάνειας ενός δείγματος τιμών $y_1,y_2,\ldots,y_n$ μιας μεταβλητής Y που ακολουθεί την κατανομή Bernoulli,με ανεξάρτητες μεταβλητές $x_i=(x_{i0}, x_{i1} \ldots x_{ik})$,$p_i$ η αντίστοιχη πιθανότητα επιτυχίας και $x_{i0}=1$ δίνεται από τον τύπο:

$$L(\beta) = \prod_{i=1}^{n} p_i^{y_i} (1 - p_i)^{1-y_i} \quad (4.5)$$

Αν θεωρήσουμε ότι $p_i = \frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}}$, δηλαδή ότι έχουμε μια ανεξάρτητη μεταβλητή, και λογαριθμήσουμε τη σχέση (4.5) θα πάρουμε:

$$\ln L(\beta) = \ln\left(\prod_{i=1}^{n} p_i^{y_i} (1 - p_i)^{1-y_i}\right) =$$

$$\sum_{i=1}^{n} [\ln(p_i^{y_i}) + \ln(1 - p_i)^{1-y_i}] =$$

$$\sum_{i=1}^{n} [y_i \ln(p_i) + (1 - y_i) \ln(1 - p_i)] =$$

Παραγωγίζοντας έχουμε :

$$\frac{\partial \ln L(\beta)}{\partial \beta_j} = \sum_{i=1}^{n} y_i x_{ij} - \sum_{i=1}^{n} x_{ij} p_i \text{ , j=0,1}$$

$$= \sum_{i=1}^{n} (y_i - p_i) x_{ij}$$

Εξισώνοντας με το μηδέν προκύπτουν οι εξισώσεις :

$$\sum_{i=1}^{n} (y_i - p_i) = 0 \quad (4.6)$$

$$\sum_{i=1}^{n} [(y_i - p_i) x_i] = 0 \quad (4.7)$$

Το παραπάνω σύστημα αποτελείται από τις μη γραμμικές εξισώσεις (4.6), (4.7) και λύνεται μόνο με επαναληπτικές μεθόδους, όπως αυτή του Newton-Raphson.Η λύση του συστήματος μας δίνει τις τιμές των εκτιμητριών $\hat{\beta}$ (Νοταρά, 2020) Οι εκτιμήτριες αυτές είναι οι τιμές που μεγιστοποιούν τη συνάρτηση πιθανοφάνειας, ή με άλλα λόγια μεγιστοποιούν την πιθανότητα να παρατηρήσουμε τα δεδομένα $y_1, y_2, y_3 \ldots y_n$.

ΚΕΦΑΛΑΙΟ 5
ΥΛΟΠΟΙΗΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ ΤΗΣ ΛΟΓΙΣΤΙΚΗΣ ΠΑΛΙΝΔΡΟΜΗΣΗΣ ΣΤΗΝ ΑΝΙΧΝΕΥΣΗ ΤΩΝ FAKE NEWS

Για την επίλυση του συγκεκριμένου προβλήματος αρχικά έπρεπε να βρούμε τα δεδομένα που θα χρησιμοποιήσουμε. Γενικά, στο συγκεκριμένο πρόβλημα είναι δύσκολο να βρεθεί μεγάλος όγκος δεδομένων, κι αυτό συμβαίνει επειδή ο χαρακτηρισμός ενός άρθρου ως fake ή real χειροκίνητα είναι μια πολύ χρονοβόρα διαδικασία , αφού χρειάζεται η εξέταση μεγάλου όγκου ειδήσεων. Ωστόσο, υπάρχει ένα σύνολο δεδομένων με το οποίο δουλεύουν αρκετοί ερευνητές και πειραματίζονται με διάφορους αλγορίθμους και αυτό χρησιμοποιήσαμε και εμείς για να τροφοδοτήσουμε τον αλγόριθμο μας με δεδομένα. Πρόκειται για δυο σύνολα δεδομένων με 23.489 περιπτώσεις άρθρων χαρακτηρισμένων ως fake news και 21.418 περιπτώσεις άρθρων χαρακτηρισμένων real news το άλλο (Bisaillon, 2019).

Στο πρόβλημα μας θεωρούμε πως ο χαρακτηρισμός αυτός των άρθρων και η ταξινόμηση τους σε fake news και real news αντίστοιχα είναι αξιόπιστος. Στόχος μας είναι , να επιλέξουμε τον κατάλληλο αλγόριθμο , ο οποίος εκπαιδευόμενος από αυτά τα δεδομένα θα μπορεί να προβλέπει εάν κάποιο άρθρο είναι real news ή fake news.

Τα ιδιαίτερα χαρακτηριστικά του προβλήματος μας , είναι ότι έχουμε να δουλέψουμε με δεδομένα κειμένου και ότι ανήκει στην κατηγορία των προβλημάτων δυαδικής ταξινόμησης, καθώς μόνο δυο δυνατά σενάρια υπάρχουν για κάθε είδηση ή να είναι αληθινή ή όχι. Επίσης, με δεδομένο ότι τα κείμενα μας είναι ήδη ταξινομημένα ως fake ή real , από τον ορισμό που δόθηκε και παραπάνω αντιλαμβανόμαστε ότι έχουμε να κάνουμε με ένα πρόβλημα επιβλεπόμενης μάθησης. Έτσι, μετά από διεξοδική αναζήτηση καταλήξαμε για τους λόγους που αναφέρθηκαν παραπάνω πως ένας αρκετά καλό μοντέλο για το πρόβλημα μας είναι αυτός της λογιστικής παλινδρόμησης.

5.1. Επεξεργασία των Δεδομένων

Τα δεδομένα κειμένου είναι ένας δύσκολος τύπος δεδομένων προς επεξεργασία. Αυτό συμβαίνει γιατί η φυσική γλώσσα είναι φτιαγμένη για να είναι κατανοητή από τους ανθρώπους αλλά όχι από τον υπολογιστή. Παράλληλα τα στατιστικά μοντέλα πρόβλεψης δέχονται σαν ορίσματα αριθμητικά δεδομένα. Έτσι, είναι απαραίτητο να αναπαραστήσουμε τα δεδομένα μας με αριθμητικό τρόπο. Η διαδικασία επεξεργασίας των δεδομένων κειμένου είναι μέρος της Επεξεργασίας Φυσικής Γλώσσας ενώ η διαδικασία αναπαράστασης του κειμένου σε αριθμούς που χρησιμοποιούμε ονομάζεται Term frequency-Inverse document frequency Vectorizer (Tf-Idf Vectorizer).

5.1.1 Καθαρισμός Δεδομένων

Η πρώτη μας εργασία στα προβλήματα αυτά αφού εισάγουμε τα δεδομένα μας , είναι αυτό που ονομάζεται καθαρισμός δεδομένων(data cleansing). Συγκεκριμένα, η μεθοδολογία που ακολουθήσαμε ήταν να αφαιρέσουμε από κάθε άρθρο τις στήλες με τον τίτλο του και την ημερομηνία του και δουλέψαμε μόνο με το κείμενο. Στο ίδιο το κείμενο ,επίσης, αφαιρέσαμε τα σημεία στίξης και μετατρέψαμε το σύνολο τους σε

πεζά για να υπάρχει μια ομοιογένεια και να μην δημιουργηθούν τυχόν προβλήματα ,όπως για παράδειγμα η ανάγνωση από τον υπολογιστή δύο ίδιων λέξεων ως διαφορετικές λόγω της χρήσης πεζών ή κεφαλαίων γραμμάτων. Τέλος, από τα κείμενα αφαιρούνται και οι λεγόμενες «stopwords» που είναι συνηθισμένες λέξεις , όπως άρθρα και προθέσεις που επαναλαμβάνονται συχνά και ενώ δεν επηρεάζουν ουσιαστικά το νόημα του κειμένου , επηρεάζουν πολύ την ανάλυση μας.

5.1.2. Term Frequency (TF)

Είναι ένας συνηθισμένος αλγόριθμος για να μετατρέψουμε το κείμενο σε μια αναπαράσταση αριθμών. Αρχικά δημιουργεί ένα λεξικό με όλες τις διαφορετικές λέξεις που υπάρχουν στα κείμενα μας. Για παράδειγμα αν έχουμε k διαφορετικές λέξεις t_k στα έγγραφα μας το αντίστοιχο λεξικό θα είχε τέτοια μορφή:

$$E(t)=\{t_1:0,\ t_2:1,\ t_3:2,\ t_4:3,\ldots,t_k:k-1\}$$

Στη συνέχεια μετατρέπει κάθε έγγραφο σε ένα διάνυσμα με διάσταση τόση όσες είναι οι λέξεις στο λεξικό μας. Κάθε στοιχείο του διανύσματος είναι η συχνότητα εμφάνισης κάθε λέξης του λεξικού που δημιουργήθηκε στο έγγραφο αυτό και η θέση του στο διάνυσμα αυτό είναι ο δείκτης της αντίστοιχης λέξης στο λεξικό. Ορίζουμε τη συνάρτηση term frequency ως εξής:

$$tf(t,d)=\sum_{x\in d} fr(x,t)$$

,όπου d το έγγραφο και $fr(x,t)$ συνάρτηση που ορίζεται ως:

$$fr(x,t)=\begin{cases}1, & \alpha\nu\ x=t\\ 0, & \delta\iota\alpha\varphi o\rho\varepsilon\tau\iota\kappa\acute{\alpha}\end{cases}$$

5.1.3. Inverse Document Frequency (IDF)

Είναι ο φυσικός λογάριθμος του αριθμού των συνολικών εγγράφων προς τον αριθμό των εγγράφων που περιέχουν τη λέξη t. Αυτή η ποσότητα είναι αντιστρόφως ανάλογη με τον αριθμό των εγγράφων που περιέχουν τη λέξη t και άρα είναι ένα μέτρο σπανιότητας της κάθε λέξης συνολικά ανάμεσα σε όλα τα κείμενα.

$$idf\,(t)=\log\frac{N}{\{d:t\in d\}}$$

,όπου $\{d:t\in d\}$ ο αριθμός των εγγράφων όπου η λέξη t εμφανίζεται και N ο αριθμός των συνολικών εγγράφων.

5.1.4. Term Frequency-Inverse document frequency (TF-IDF)

Ο τύπος που υπολογίζει τελικά το tf-idf είναι :

$$\text{tf-idf}(t)=\text{tf}(t,d)\times\ \text{idf}(t)$$

Έτσι η ποσότητα αυτή αυξάνεται όταν έχουμε μια υψηλή συχνότητα για μια λέξη σε ένα έγγραφο (τοπική παράμετρος) και χαμηλή συχνότητα εμφάνισης της λέξης σε όλα τα έγγραφα (παγκόσμια παράμετρος).Έτσι, το διάνυσμα ενός εγγράφου d γίνεται :

$$V_d=\left(tf-idf(t_1),tf-idf(t_2),tf-idf(t_3),....,tf-idf(t_k)\right)$$

Τελικά , αν θεωρήσουμε ότι έχουμε N έγγραφα συνολικά δημιουργείται ένας πίνακας

Nxk διαστάσεων:

$$M_{Nxk}=\begin{bmatrix} tf-idf(t1)(1) & \cdots & tf-idf(tk)(1) \\ \vdots & \ddots & \vdots \\ tf-idf(t1)(N) & \cdots & tf-idf(tk)(N) \end{bmatrix}$$

5.2. Εφαρμογή του Μοντέλου

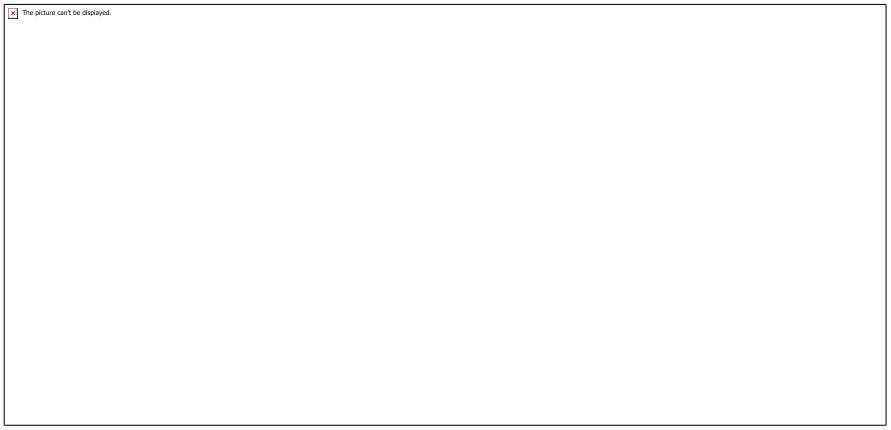
Η βασική ιδέα σε ένα πρόβλημα επιβλεπόμενης μάθησης είναι η εκπαίδευση του αλγορίθμου από μέρος των δεδομένων έτσι ώστε να μπορεί στη συνέχεια να προβλέπει το αποτέλεσμα όταν η εισαγωγή είναι νέα δεδομένα. Έτσι τα δεδομένα χωρίζονται σε δυο μέρη τα train data και τα test data συνήθως σε μια αναλογία κοντά στο 80-20 διότι όσο μεγαλύτερος είναι ο αριθμός των δεδομένων που χρησιμοποιούμε για να εκπαιδεύσουμε τον αλγόριθμο μας τόσο πιο ακριβής θα είναι η προβλεπτική του ικανότητα. Τα test data είναι αυτά που χρησιμοποιούμε αφού εκπαιδεύσουμε το μοντέλο για να ελέγξουμε πόσο ακριβείς είναι οι προβλέψεις του.

Συγκεκριμένα , στο πρόβλημα μας αφού χωρίσαμε το σύνολο των άρθρων μας σε train και test data τροφοδοτήσαμε το μοντέλο της λογιστικής παλινδρόμησης με τα train data. Η ανεξάρτητη μεταβλητή X είναι ο πίνακας $M_{N\times k}$ που περιέχει τα tf-idf των k λέξεων στα N άρθρα που χρησιμοποιούμε ως δεδομένα εκπαίδευσης. Η εξαρτημένη μεταβλητή Y είναι το διάνυσμα με τις ετικέτες real ή fake(δηλαδή 0 ή 1) για κάθε ένα από τα N άρθρα. Δηλαδή :

$$X=\begin{bmatrix} x_{11} & x_{12} & . & . & . & x_{1k} \\ x_{21} & . & & & & . \\ . & & . & & & . \\ . & & & . & & . \\ . & & & & . & . \\ x_{n1} & . & . & . & . & x_{nk} \end{bmatrix}_{n\times k}, Y=\begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_n \end{bmatrix}_{n\times 1}$$

Στη συνέχεια μέσω της μεθόδου εκτίμησης μέγιστης πιθανοφάνειας υπολογίζονται οι παράμετροι b και βρίσκουμε τη σιγμοειδή συνάρτηση:

$$\sigma(z)=\frac{1}{1+e^{-z}}\text{ , όπου } z=e^{b_0+b_1X}.$$



Γράφημα 2: Αναπαράσταση στο δισδιάστατο χώρο των σημείων (x,y) και της σιγμοειδούς συνάρτησης παλινδρόμησης. Πηγή: https://realpython.com/logistic-regression-python/?fbclid=IwAR0b8sKdRI8hX7bXkPQvnO1RsQlU_nFeUboBmAKszfgDmOscmwG2QPMuX3E

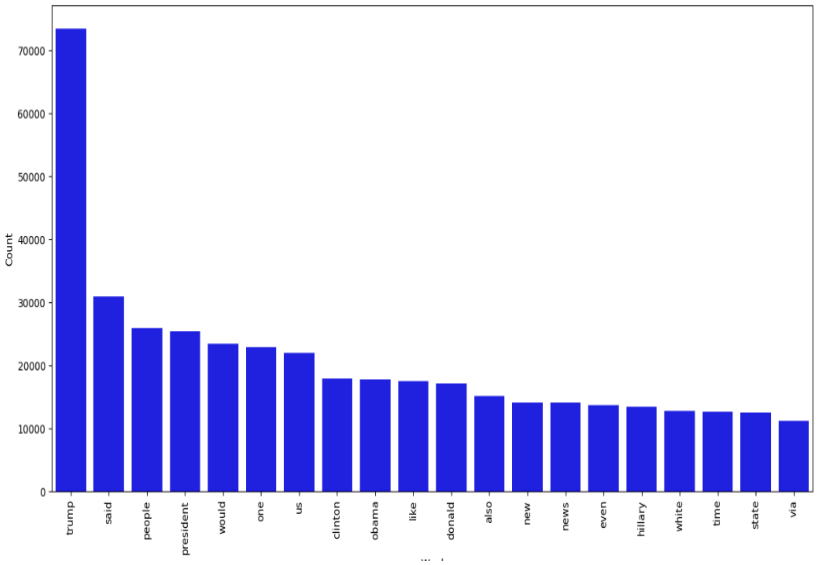
Έτσι, έχοντας πλέον τη συνάρτηση παλινδρόμησης μπορούμε να την τροφοδοτούμε με δεδομένα Χ και μέσω της συνάρτησης να παίρνουμε μια πρόβλεψη για την πιθανότητα η μεταβλτή Υ να ανήκει σε μια από τις δύο κατηγορίες. Συγκεκριμένα, για P_0=0.5 κατώφλι ταξινόμησης, αν P(x)>P_0 τότε y=1 δηλαδή το μοντέλο ταξινομεί το άρθρο ως fake news, ενώ αν P(x)<P_0 τότε y=0, δηλαδή ταξινομεί το άρθρο ως real news. Τέλος, με τα test data ελέγχουμε τις επιδόσεις του μοντέλου μας δηλαδή σε τι ποσοστό οι προβλέψεις μας είναι σωστές.

Για την υλοποίηση των παραπάνω χρησιμοποιήθηκε Γλώσσα Python.Ο λόγος που επιλέχθηκε η συγκεκριμένη γλώσσα είναι η απλότητα της , η υψηλή αναγνωσιμότητα της αλλά κυρίως τα εργαλεία που προσφέρει. Συγκεκριμένα, παρέχει διάφορα πακέτα και βιβλιοθήκες τα οποία είναι ιδιαίτερα χρήσιμα στην ανάλυση και στην επεξεργασία δεδομένων καθώς και στην επίλυση προβλημάτων που απαιτούν τεχνικές μηχανικής μάθησης και μαθηματικών στατιστικών μοντέλων. Η υλοποίηση του κώδικα παρατίθεται στο Παράρτημα.

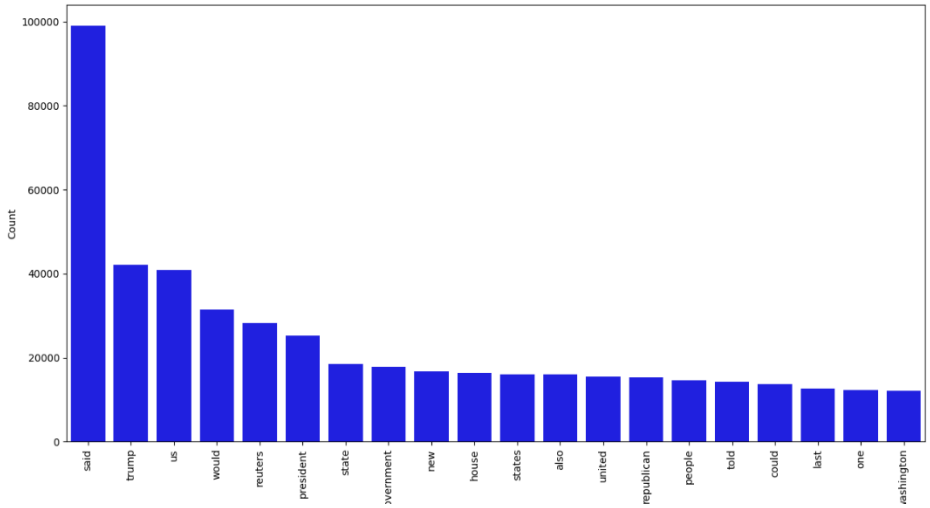
ΚΕΦΑΛΑΙΟ 6
ΣΥΜΠΕΡΑΣΜΑΤΑ

Σε αυτό το κεφάλαιο θα παρουσιάσουμε τα αποτελέσματα που παίρνουμε όταν εκτελούμε τον προαναφερθέντα κώδικα καθώς και σχετικά συμπεράσματα που προκύπτουν από την έρευνα.

Αρχικά εμφανίζεται ένα ραβδόγραμμα το οποίο παρουσιάζει τις 20 πιο συχνά εμφανιζόμενες λέξεις μέσα στα κείμενα μας τα οποία είναι ταξινομημένα ως fake.

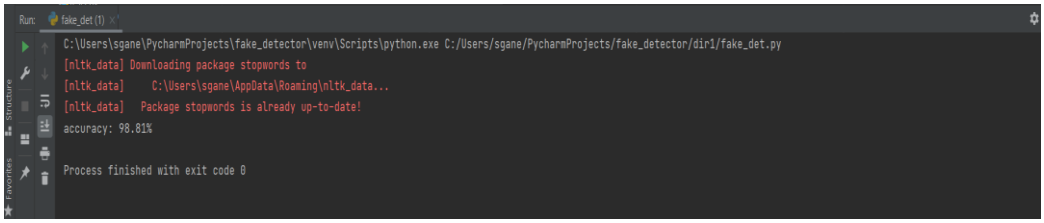


Στη συνέχεια εμφανίζεται ένα δεύτερο ραβδόγραμμα, το οποίο αυτή τη φορά παρουσιάζει τις 20 πιο συχνά εμφανιζόμενες λέξεις στα κείμενα μας που είναι ταξινομημένα ως real.



Η εμφάνιση αυτών των δύο γραφημάτων δεν στοχεύει τόσο στην αποκόμιση σημαντικών συμπερασμάτων όσον αφορά τα δεδομένα μας και την ανάλυση τους αλλά

μας δίνει τη δυνατότητα να εξερευνήσουμε τα δεδομένα μας και πιθανώς να εξάγουμε κάποιο ποιοτικό στοιχείο από αυτά. Συγκρίνοντας τα, λοιπόν, έχει μια αξία να παρατηρήσουμε ότι στο δεύτερο γράφημα η λέξη «reuters» που αναφέρεται στο διεθνές πρακτορείο ειδήσεων καταλαμβάνει την πέμπτη θέση ενώ αντίθετα στο πρώτο γράφημα δεν εμφανίζεται καθόλου μέσα στις είκοσι πιο συχνές λέξεις. Άρα η πεποίθηση μας ότι τα fake news βασίζονται κυρίως σε ισχυρισμούς που δεν έχουν τόσο στέρεα βάση όσο οι ισχυρισμοί ενός μεγάλου και έγκυρου διεθνούς πρακτορείου ειδήσεων, σε κάποιο βαθμό επιβεβαιώνεται.



Τέλος έχουμε ως αποτέλεσμα εμφάνισης του κώδικα το ποσοστό ακρίβειας του μοντέλου μας. Όπως εξηγήθηκε και νωρίτερα το μοντέλο μας εκπαιδεύτηκε από το 80% των δεδομένων που το τροφοδοτήσαμε και έλεγξε τις επιδόσεις του στο υπόλοιπο 20%. Παρατηρούμε ότι οι προβλέψεις μας είναι σωστές κατά 98,81%.

Συμπερασματικά, λοιπόν, μπορούμε να πούμε ότι το μοντέλο που υλοποιήσαμε είναι εύκολα υλοποιήσιμο από άποψη κώδικα και αρκετά αποδοτικό άρα μπορεί να αποτελέσει τη βάση για περαιτέρω έρευνα στον τομέα της ανίχνευσης των Fake News. Συγκεκριμένα, λαμβάνοντας υπόψη και τις πιο πρόσφατες τεχνολογικές εξελίξεις στον τομέα της αυτόματης παραγωγής κειμένου μπορεί να δοκιμαστεί επίσης σε κείμενα που έχουν παραχθεί από μηχανές και να διαπιστωθεί αν είναι εξίσου αποδοτικό. Παράλληλα, θα μπορούσε να γίνει η αφορμή για την αντιμετώπιση παρόμοιων φαινομένων διάδοσης Fake News και στη χώρα μας εμπνέοντας ερευνητές να δημιουργήσουν ένα σύνολο δεδομένων αποτελούμενο από ελληνικά κείμενα και τροποποιώντας κατάλληλα τις βιβλιοθήκες της Python έτσι ώστε να μπορεί να εφαρμοστεί και για κείμενα γραμμένα με ελληνικούς χαρακτήρες. Κάτι τέτοιο θα μπορούσε να αποτελέσει ένα σημαντικό εργαλείο για κάθε επιχείρηση, οργανισμό , τα ΜΚΔ ακόμα και για τις κυβερνήσεις στην προσπάθεια τους να υλοποιήσουν μια πολιτική προστασίας του κοινωνικού συνόλου από τα Fake News.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Allcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives*. doi:10.1257/jep.31.2.211

Ambalina, L. (2020). *Tools to Spot Deepfakes and AI-Generated Text*. Ανάκτηση από KDnuggets: https://www.kdnuggets.com/2020/06/dont-click-this-how-spot-deepfakes.html

Benjamin, G., & Gilis, A. (2021). *Turing Test*. Ανάκτηση από SearchEnterpriseAI: https://searchenterpriseai.techtarget.com/definition/Turing-test

Bisaillon, C. (2019). *Fake and real news dataset*. Ανάκτηση από Kaggle: https://www.kaggle.com/clmentbisaillon/fake-and-real-news-dataset

Brody, D., & Meier, D. (2018). *How to model fake news.*

Brownell, K. D., & Warner, K. E. (2009). The Perils of Ignoring History: Big Tobacco Played Dirty and Millions Died. How Similar Is Big Food? *Milbank Quarterly*. doi:10.1111/j.1468-0009.2009.00555.x

Brownlee, J. (2016). *Machine Learning Mastery*. Ανάκτηση από Logistic Regression for Machine Learning: https://machinelearningmastery.com/logistic-regression-for-machine-learning/

Cassauwers, T. (2019). Can artificial intelligence help end fake news? *Horizon*. Ανάκτηση από https://ec.europa.eu/research-and-innovation/en/horizon-magazine/can-artificial-intelligence-help-end-fake-news

Chaudhary, M. (2020). *TF-IDF Vectorizer scikit-learn*. Ανάκτηση από Medium: https://medium.com/@cmukesh8688/tf-idf-vectorizer-scikit-learn-dbc0244a911a

Corner, J. (2017). Fake news,post truth and media-political change. *Media,Culture and Society*. doi:https://doi.org/10.1177/0163443717726743

Dickson, E. (2019). *Deepfake Porn Is Still a Threat, Particularly for K-Pop Stars*. Ανάκτηση από RollingStone: https://www.rollingstone.com/culture/culture-news/deepfakes-nonconsensual-porn-study-kpop-895605/

EAVI. (2017). *EAVI.EU*. Ανάκτηση από Infographic:Beyond Fake News- 10 types of Misleading News: https://eavi.eu/beyond-fake-news-10-types-misleading-info/

European Comission. (2018). *A multi-dimensional approach to disinformation.* EU. doi:10.2759/0156

Fagni, T., Falchi, F., Gambini, M., Martella, A., & Tesconi, M. (2021). TweepFake: About detecting deepfake tweets. *PLoS One*. doi:10.1371/journal.pone.0251415

Fawzi, N. (2017). Beyond policg agenda-setting: political actors' and journalists' perceptions of news media influence across all stages of the political process. *Information Communication and Society*. doi:https://doi.org/10.1080/1369118X.2017.1301524

Foote, K. D. (2019). *A Brief History of Natural Language Processing (NLP)*. Ανάκτηση από Dataversity: https://www.dataversity.net/a-brief-history-of-natural-language-processing-nlp/

Fourney, A. (2017). Geographic and temporal trends in fake news consumption during the 2016 US presidential election. *International Conference on Information and Knowledge Management, Proceedings.* Association for Computing Machinery. doi:10.1145/3132847.3133147

Funk, C. (2017). *Mixed Messages about Public Trust in Science*. Ανάκτηση από Pew Research Center: https://www.pewresearch.org/science/2017/12/08/mixed-messages-about-public-trust-in-science/

Graham, D. (2019). Some Real News About Fake News. *The Atlantic*. Ανάκτηση από https://www.theatlantic.com/ideas/archive/2019/06/fake-news-republicans-democrats/591211/

Hao, K., & Heaven, W. D. (2020). *The year deepfakes went mainstream*. Ανάκτηση από MIT Technology Review: https://www.technologyreview.com/2020/12/24/1015380/best-ai-deepfakes-of-2020/

Harsin, J. (2018). A CRITICAL GUIDE TO FAKE NEWS: FROM COMEDY TO TRAGEDY. *Pouvoirs: Revue d'Etudes Constitutionnelles et Politiques*.

Hern, A. (2019). New AI fake text generator may be too dangerous to release, say creators. *The Guardian*. Ανάκτηση από https://www.theguardian.com/technology/2019/feb/14/elon-musk-backed-ai-writes-convincing-news-fiction

Hviid, A., Hansen, V. J., Frisch, M., & Melbye, M. (2019). Measles, Mumps, Rubella Vaccination and Autism: A Nationwide Cohort Study. *Annals of Internal Medicine*. doi:10.7326/M18-2101

Ireton, C., Posettie, J., & Unesco. (2018). *Journalism, fake news & disinformation : handbook for journalism education and training / Cherilyn Ireton and Julie Posetti.*

Kestenbaum, L. A., & Feemster, K. A. (2015). Identifying and addressing vaccine hesitancy. *Pediatric Annals*. doi:10.3928/00904481-20150410-07

Larson, J. H. (2018). The biggest pandemic risk? Viral misinformation. *Nature*. doi:https://doi.org/10.1038/d41586-018-07034-4

Lazer, D., Baum, M., Benchler, Y., Berinsky, A., Greenhill, K., Menczer, F., . . . Zittrain, J. (2018). The science of fake news. *PubMed*. doi:10.1126/science.aao2998

Lee, T. (2019). The global rise of “fake news” and the threat to democratic elections in the USA. *Public Administration and Policy: An Asia-Pacific Journal*. doi:https://doi.org/10.1108/PAP-04-2019-0008

Li, Y., & Lyu, S. (2018). Exposing DeepFake Videos By Detecting Face Warping Artifacts. *Conference on Computer Vision and Pattern Recognition.*

Loos, E., & Nijenhuis, J. (2020). Consuming Fake News: A Matter of Age? The Perception of Political Fake News Stories in Facebook Ads. Στο H. A. Society (Επιμ.), *6th International Conference, ITAP 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part III*. Springer International Publishing.

Lustig, C., Pine, K., Nardie, B., Irani, L., Lee, M. K., Nafus, D., & Sandvig, C. (2016). Algorithmic authority: The ethics, politics, and economics of algorithms that interpret, decide, and manage. *34th Annual CHI Conference on Human Factors in Computing Systems, CHI EA 2016 - San Jose, United States.* doi:10.1145/2851581.2886426

Lyu, S. (2018). Detecting ‘deepfake’ videos in the blink of an eye. *The Conversation*. Ανάκτηση από https://theconversation.com/detecting-deepfake-videos-in-the-blink-of-an-eye-101072

Lyu, S. (2019). Detecting deepfakes by looking closely reveals a way to protect against them. *The Conversation*. Ανάκτηση από https://theconversation.com/detecting-deepfakes-by-looking-closely-reveals-a-way-to-protect-against-them-119218

Maklin, C. (2019). *TF IDF | TFIDF Python Example*. Ανάκτηση από Towards Data Science: https://towardsdatascience.com/natural-language-processing-feature-engineering-using-tf-idf-e8b9d00e7e76

Manning, C., Raghavan, P., & Shutze, H. (2008). *Retrieval, Introduction to Information.*

Menard, S. (2010). *Logistic Regression: From Introductory to Advanced Concepts and Applications.* London: Thousand Oaks.

Obada, R. (2019). Sharing Fake News about Brands on Social Media: a New Conceptual Model Based on Flow Theory. *Argumentum: Journal of the Seminar of Discursive Logic, Argumentation Theory and Rhetoric*.

Pant, A. (2019). *Introduction to Logistic Regression*. Ανάκτηση από Towards Data Science: https://towardsdatascience.com/introduction-to-logistic-regression-66248243c148

Rao, T., & Andrade, C. (2011). The MMR vaccine and autism: Sensation, refutation, retraction, and fraud. *Indian Journal of Psychiatry*. doi:10.4103/0019-5545.82529

Sample, I. (2020). What are deepfakes – and how can you spot them? *The Guardian*. Ανάκτηση από https://www.theguardian.com/technology/2020/jan/13/what-are-deepfakes-and-how-can-you-spot-them

Smith, J., Thompson, S., & Lee, K. (2016). Death and taxes: The framing of the causes and policy responses to the illicit tobacco trade in Canadian newspapers. *Cogent Social Sciences*. doi:https://doi.org/10.1080/23311886.2017.1325054

Stemwedel, J. D. (2011). Drawing the line between science and pseudo-science. *Scientific American*.

Stojiljković , M. (Logistic Regression in Python). *Logistic Regression in Python*. Ανάκτηση από Real Python: https://realpython.com/logistic-regression-python/?fbclid=IwAR0b8sKdRI8hX7bXkPQvnO1RsQlU_nFeUboBmAKszfgDmOscmwG2QPMuX3E#reader-comments

Subasi, C. (2019). *LOGISTIC REGRESSION CLASSIFIER*. Ανάκτηση από Towards Data Science: https://towardsdatascience.com/logistic-regression-classifier-8583e0c3cf9

Tandoc, E. (2017). Defining "Fake News": A typology of scholarly definitions. *Digital Journalism*. doi:10.1080/21670811.2017.1360143

Watts, J. (2018). We have 12 years to limit climate change catastrophe, warns UN. *The Guardian*.

WHO. (2018). *Measles cases hit record high in the European Region*. Ανάκτηση από Word Health Organization: https://www.euro.who.int/en/media-centre/sections/press-releases/2018/measles-cases-hit-record-high-in-the-european-region

Wooley, S. C., & Howard, P. N. (2017). *Computational Propaganda Worldwide:*. Oxford, UK: Project on Computational Propaganda.

Xenopoulos, P. (2017). *Why is machine learning happening now?* Ανάκτηση από Medium: https://medium.com/@peterx/machine-learning-why-is-everyone-doing-it-now-98b0ae6e3fc

Yang, X., Li, Y., & Lyu, S. (2018). Exposing Deep Fakes Using Inconsistent Head Poses.

Yuezun, L., Ming-Ching, C., & Siwei, L. (2018). In Ictu Oculi: Exposing AI Created Fake Videos by Detecting Eye Blinking. *IEEE International Workshop on Information Forensics and Security (WIFS)*, 1-7. doi:10.1109/WIFS.2018.8630787

Zellers, R., Holtzman, A., Rashkin, H., Bisk, Y., Farhadi, A., Roesner, F., & Choi, Y. (2019). Defending Against Neural Fake News. *Conference on Neural Information Processing Systems.*

Ευρωπαϊκό Κοινοβούλιο. (2021). *Τι είναι η τεχνητή νοημοσύνη και πώς χρησιμοποιείται;*. Ανάκτηση από European Parliament: https://www.europarl.europa.eu/news/el/headlines/society/20200827STO85804/ti-einai-i-techniti-noimosuni-kai-pos-chrisimopoieitai

Ευτυχίου, Α. (2019). *Αλγόριθμοι μηχανικής μάθησης και εφαρμογές σε ιατροβιολογικά προβλήματα.* Ανάκτηση από https://nemertes.library.upatras.gr/jspui/bitstream/10889/13208/1/eytychia_anastasiou_diplomatiki.pdf

Εφημερίδα των Συντακτών. (2019). «Καθαρίζει» με 500.000 λίρες η Facebook για το σκάνδαλο της Cambridge Analytica. Ανάκτηση από Efsyn: https://www.efsyn.gr/oikonomia/diethnis-oikonomia/216843_katharizei-me-500000-lires-i-facebook-gia-skandalo-tis

Θεοδωρακόπουλος, Π. (2006). *Προπαγάνδα η ένδοξη.* Σιδέρης Ι.

Καρώνη, Χ., & Οικονόμου, Π. (2010). *Στατιστικά Μοντέλα Παλινδρόμησης.* Συμεών.

Κούτρας, Μ. (2011). Ανάλυση Παλινδρόμησης. Ανάκτηση από http://www.unipi.gr/faculty/mkoutras/regres/regres1_1.pdf

Νοταρά, Σ. (2020). *Το Λογιστικό Μοντέλο Παλινδρόμησης και Δέντρα Ταξινόμησης.* Αθήνα.

Παπαθανασόπουλος, Σ. (2017). Τα σύγχρονα Μέσα και η πολιτική επικοινωνία. *Ελληνική Επιθεώρηση Πολιτικής Επιστήμης*. doi:https://doi.org/10.12681/hpsa.15184

Παπαϊωάννου, Γ. (2018). *Τι κρύβει το σκάνδαλο της Cambridge Analytica και του Facebook.* Ανάκτηση από Lifo: https://www.lifo.gr/now/tech-science/ti-krybei-skandalo-tis-cambridge-analytica-kai-toy-facebook

Πλειός, Γ. (2018). Fake News: ψευδεπίγραφες ειδήσεις. Εξαίρεση ή κανόνας ;. *Η Αυγή*.

Πλειός, Γ. (2019). *Fake News: Τα 4+1 βασικά γνωρίσματα.* Ανάκτηση από tvxs: https://tvxs.gr/news/egrapsan-eipan/fake-news-ta-41-basika-gnorismata

Πουλακιδάκος, Σ. (2013). *Η προπαγάνδα ως θεμελιώδες συστατικό του δημόσιου λόγου: η παρουσίαση του «μνημονίου» από τα ελληνικά ΜΜΕ.*

Φουσκάκης, Δ. (2013). *Ανάλυση Δεδομένων με Χρήση της R.* Τσότρας.

ΠΑΡΑΡΤΗΜΑ

Παρακάτω παρατίθεται ο κώδικας που υλοποιήθηκε :

```
import pandas as pd
fake = pd.read_csv("/Users/sgane/Documents/Διπλωματικη/Fake.csv")
true = pd.read_csv("/Users/sgane/Documents/Διπλωματικη/True.csv")
fake['target'] = 'fake'
true['target'] = 'real'
data = pd.concat([fake, true]).reset_index(drop=True)

# shuffle data
from sklearn.utils import shuffle
data = shuffle(data)
data = data.reset_index(drop=True)

# data cleansing
data.drop(["title"], axis=1, inplace=True)
data.drop(["date"], axis=1, inplace=True)
data.drop["subject"],axis=1,inplace=True)
data['text'] = data['text'].apply(lambda x: x.lower())
import string



def punctuation_removal(text):
all_list = [char for char in text if char not in string.punctuation]
clean_str = ''.join(all_list)
return clean_str
```

```
data['text'] = data['text'].apply(punctuation_removal)

import nltk



nltk.download('stopwords')

from nltk.corpus import stopwords



stop = stopwords.words('english')

data['text'] = data['text'].apply(lambda x: ' '.join(word for word in x.split() if word not in stop))



from nltk import tokenize

import seaborn as sns

import matplotlib.pyplot as plt

token_space = tokenize.WhitespaceTokenizer()

def counter(text,column_text,quantity):

all_words=' '.join([text for text in text[column_text]])

token_phrase=token_space.tokenize(all_words)

frequency=nltk.FreqDist(token_phrase)

df_frequency = pd.DataFrame({"Word": list(frequency.keys()),"Frequency": list(frequency.values())})

df_frequency = df_frequency.nlargest(columns="Frequency", n=quantity)

plt.figure(figsize=(12, 8))

ax = sns.barplot(data=df_frequency, x="Word", y="Frequency", color='blue')

ax.set(ylabel="Count")

plt.xticks(rotation='vertical')

plt.show()

counter(data[data["target"]=="fake"],"text",20)
```

```
counter(data[data["target"]=="real"],"text",20)

from sklearn.model_selection import train_test_split

from sklearn.pipeline import Pipeline

from sklearn.feature_extraction.text import CountVectorizer

from sklearn.feature_extraction.text import TfidfTransformer

from sklearn.metrics import accuracy_score


X_train,X_test,y_train,y_test = train_test_split(data['text'], data.target, test_size=0.2,
random_state=42)




# Vectorizing and applying TF-IDF

from sklearn.linear_model import LogisticRegression

pipe = Pipeline([('vect', CountVectorizer()),

('tfidf', TfidfTransformer()),

('model', LogisticRegression())])

# Fitting the model

model = pipe.fit(X_train, y_train)

# Accuracy

prediction = model.predict(X_test)

print("accuracy: {}%".format(round(accuracy_score(y_test, prediction)*100,2)
```