

# Τεχνητή Νοημοσύνη

2<sup>η</sup> Εργασία

Ευράφης Γιώργος  
Α.Μ.:2022201800138



05/2024

# Περιεχόμενα

Δίκτυα Bayes.....	3
2.1.....	3
2.2.....	4
Φίλτρο Ανίχνευσης Spam Μηνυμάτων με χρήση αλγορίθμου Naive Bayes.....	5
3.2-3.3.....	5

# Δίκτυα Bayes

## 2.1

Το δίκτυο Bayes έχει κατασκευαστεί όπως ορίζεται στην εκφώνηση. Συγκεκριμένα έχουμε τους 6 κόμβους που ζητούνται και με βάση τις πληροφορίες που μας δίνονται για την συσχέτιση των κόμβων και τα ποσοστά του να συμβεί κάτι δεδομένου κάτι άλλο, έχουμε αποφασίσει ποιοι θα είναι οι γονείς του κάθε κόμβου καθώς και ο πίνακας πιθανοτήτων του.

Όσον αφορά τον κόμβο “Taksidevei”, με τις πληροφορίες που μας δίνονται βλέπουμε πως δεν έχει κάποιον γονιό και έχει 5% πιθανότητα να συμβαίνει συναλλαγή όσο κάποιος ταξιδεύει, άρα 0.05 σε τιμή. Αντίστοιχα για τον κόμβο “DiatheteiIpologisti” με τις πληροφορίες που μας δίνονται βλέπουμε πως δεν έχει κάποιον γονιό και έχει 60% πιθανότητα να συμβαίνει, άρα 0.6.

Οι κόμβοι “Apati” και “AgoraSxetikiMeIpologisti” ακολουθούν αντίστοιχη φιλοσοφία ο ένας με τον άλλον. Από τις πληροφορίες που έχουμε, μαθαίνουμε πως ο κόμβος “Apati” επηρεάζεται μόνο από το αν ο κάτοχος βρίσκεται στο εξωτερικό ή όχι, συνεπώς έχει ως γονιό τον κόμβο “Taksidevei”. Συγκεκριμένα “το 1% των συναλλαγών είναι απάτες όταν ο/η κάτοχος της κάρτας ταξιδεύει, ενώ μόνο το 0.4% των συναλλαγών είναι απάτες όταν δεν ταξιδεύει.”. Άρα κατά συνέπεια, δεδομένου του “Taksidevei” ως True ο κόμβος “Apati” λαμβάνει τιμή 0.01 ενώ όταν ο γονιός είναι False λαμβάνει τιμή 0.004. Αντίστοιχα ο κόμβος “AgoraSxetikiMeIpologisti” έχει γονιό τον κόμβο “DiatheteiIpologisti” και τιμές 0.1 (10%) για όταν ο γονιός είναι True και 0.001 (0.1%) όταν ο γονιός είναι False.

Τέλος αντίστοιχη φιλοσοφία μεταξύ τους ακολουθούν και οι κόμβοι “AgoraEksoterikou” και “AgoraDiadiktou”. Από τις πληροφορίες που μας δίνονται βλέπουμε πως ο κόμβος “AgoraEksoterikou” επηρεάζεται από τους κόμβους “Taksidevei” και “Apati”. Συνεπώς και οι δύο κόμβοι θα δηλωθούν ως γονείς του και τα ποσοστά θα δοθούν σε μορφή πίνακα που εξετάζει τις τιμές των δύο γονιών του. Έτσι όταν το “Taksidevei” είναι True, από τις πληροφορίες μας βλέπουμε πως η τιμή, ανεξάρτητα της απάτης πάει στο 90%. Ενώ όταν είναι false εξαρτάται από την “Apati” με τα αντίστοιχα ποσοστά. Ο κόμβος “AgoraDiadiktou” αντίστοιχα βλέπουμε πως έχει δύο γονείς, τον “DiatheteiIpologisti” και τον “Apati” και τα ποσοστά μας δίνονται αντίστοιχα σε πίνακα που εξετάζει όλους τους πιθανούς συνδυασμούς True – False των γονιών.

Αξίζει να σημειωθεί πως το δίκτυο Bayes μας δείχνει απλά συσχέτιση και όχι κάποια σχέση αιτίας μεταξύ του κάθε κόμβου και των γονιών του. Επίσης, για να είναι σωστό, πρέπει να μην περιέχει κύκλους.

## 2.2

Παρατηρούμε πως η πιθανότητα απάτης μειώνεται σημαντικά και πλησιάζει το 0 όταν προσθέσουμε την μαρτυρία πως η γυναίκα βρίσκεται σε ένα επαγγελματικό ταξίδι. Αυτό είναι κάτι που το περιμέναμε, καθώς η αγορά έγινε στο εξωτερικό και όχι από ονλάιν μέσα, αλλά δια ζώσης, άρα το γεγονός πως η γυναίκα βρίσκεται σε ταξίδι κάνει σημαντικά πιθανότερο το ενδεχόμενο να έκανε η ίδια την αγορά και να μην πρόκειται για απάτη.

# Φίλτρο Ανίχνευσης Spam Μηνυμάτων με χρήση αλγορίθμου Naive Bayes

## 3.2-3.3

Βλέπουμε πως ο αλγόριθμος έχει ένα ποσοστό ακρίβειας της τάξεως του 86% (0.866), το οποίο είναι αρκετά υψηλό και ικανοποιητικό, πιθανότατα καλύτερο απ' ό,τι θα ήταν αν απλά είχαμε ορισμένους περιορισμένους και συγκεκριμένους κανόνες για να κάνουν την ίδια δουλειά, αλλά αυτό βέβαια εξαρτάται και από το μέγεθος και την ποιότητα του συνόλου εκπαίδευσης. Με την εξομάλυνση Laplace το ποσοστό αυτό αυξάνεται και φτάνει το 92%, γεγονός που είναι λογικό αν λάβουμε υπόψιν το ότι με την εξομάλυνση Laplace διάφορες λέξεις που δεν ήταν στο training set και απλά αγνοήθηκαν, τώρα αξιοποιούνται.