

Νευρωνικά δίκτυα βαθιάς μάθησης για τη μη επιβλεπόμενη ανίχνευση ανωμαλιών σε δεδομένα

**Παρασκευόυπουλος Παναγιώτης**

Επιβλέπων καθηγητής: **Ιωάννης Μπούταλης**

**Περίληψη**

Περιεχόμενα

[1.Εισαγωγή 4](#_Toc68195459)

[Ανίχνευση ανωμαλιών 4](#_Toc68195460)

[Ιστορική αναδρομή 4](#_Toc68195461)

# 1.Εισαγωγή

## Ανίχνευση ανωμαλιών

Η ανίχνευση ανωμαλιών είναι η αναγνώριση έκτροπων δεδομένων ή ακραίων τιμών (outliers) σε ένα σύνολο δεδομένων. Οι ανωμαλίες αυτές μπορεί να εμφανίζονται σπάνια μέσα στα δεδομένα ή τα χαρακτηριστικά τους να διαφέρουν αρκετά από το μέσο όρο των υπολοίπων δεδομένων.

Η ανίχνευση ανωμαλιών χωρίζεται σε τρείς κατηγορίες με βάση την τεχνική εκμάθησης των μοντέλων που χρησιμοποιούμε:

Την επιβλεπόμενη ανίχνευση ανωμαλιών κατα την οποία το σύνολο των δεδομένων μας περιέχει ετικέτες που υποδεικνύουν ποιά δεδομένα είναι κανονικά (inliers) και ποιά είναι έκτροπα (outliers) και η οποία αντιμετωπίζεται σαν πρόβλημα ταξινόμισης (classification).

Την ημι-επιβλεπόμενη ανίχνευση ανωμαλιών κατα την οποία θεωρούμε ότι ένα σύνολο δεδομένων είναι κανονικά και εκπαιδεύουμε το μοντέλο μας πάνω σε αυτά και έπειτα το μοντέλο μας εμφανίζει μεγάλη απόκλιση στα έκτροπα δεδομένα στη φάση της δοκιμής (test).

Την μη επιβλεπόμενη ανίχνευση ανωμαλιών κατα την οποία δεν έχουμε καθόλου ετικέτες στα δεδομένα και το μοντέλο μας προσπαθεί να τα κατηγοριοποιήσει έχοντας ως υπόθεση ότι τα έκτροπα δεδομένα είναι πολύ λιγότερα από τα κανονικά.

## Ιστορική αναδρομή

Η χρήση τεχνικών ανίχνευση ανωμαλιών είχε προταθεί ως ανίχνευση εισβολής (intrusion detection) από την Dorothy Denning το 1986

<https://en.wikipedia.org/wiki/Anomaly_detection>

## Σύγχρονες εφαρμογές

Η ανίχνευση ανωμαλιών εφαμόζεται σε ένα ευρύ φάσμα πεδίων όπως η ανίχνευση εισβολής ή κακής χρήσης σε δίκτυα υπολογιστών, η ανίχνευση απάτης σε τραπεζικές συναλλαγές, η παρακολούθηση σωστής λειτουργίας κάποιου συστήματος, η ανίχνευση συμβάντων σε δίκτυα αισθητήρων ή η ανίχνευση διαταρραχών σε ένα οικοσύστημα και η προεπεξεργασία δεδομένων με σκοπό την αφαίρεση έκτροπων δεδομένων πριν χρησιμοποιηθούν από κάποιο άλλο μοντέλο μηχανικής μάθησης ωστέ να αυξηθεί η τελική του απόδοση.

https://en.wikipedia.org/wiki/Anomaly\_detection

## Νευρωνικά δίκτυα βαθιάς μάθησης

<http://neuralnetworksanddeeplearning.com/chap6.html>

Τα νευρωνικά δίκτυα βαθιάς μάθησης αποτελούν κομμάτι της μηχανικής μάθησης και βασίζονται στα τεχνιτά νευρωνικά δίκτυα. Τα τεχνικά νευρωνικά δίκτυα αποτελούνται συνήθως από ένα στρώμα εισαγωγής, ένα κρυφό στρώμα και ένα στρώμα εξόδου. Τα νευρωνικά δίκτυα βαθιάς μάθησης επεκτείνουν τα κρυφά στρώματα σε περισσότερα του ενός, πετυχαίνοντας έτσι καλύτερη πρακτική απόδοση από τα τεχνιτά νευρωνικά δίκτυα ειδικότερα σε μεγάλα σύνολα δεδομένων καθώς και σε σύνθετα προβλήματα όπως η αναγνώριση εικόνας και ήχου. Στην συγκεκριμένη εργασία χρησιμοποιείται ένας συγκεκριμένος τύπους νευρωνικού δικτύου βαθιάς μάθησης που ονομάζεται Αυτοκωδικοποιητής (Autoencoder).

Ο Αυτοκωδικοποιητής είναι ένα είδος νευρωνικού δικτύου το οποίο εκπαιδεύεται ώστε να αντιγράφει τα δεδομένα που δέχεται στην είσοδο, στην έξοδό του.

«»» μαθηματικά»»»» εικόνα «»»»»»

Συγκεκριμένα αποτελείται απο δύο τμήματα: τον κωδικοποιητή (encoder) ο οποίος κωδικοποιεί την είσοδο σε ένα χώρο προβολής (latent space) και τον αποκωδικοποιητή ο οποίος προσπαθεί να ανακατασκευάσει την είδοδο απο τον χώρο προβολής. Ο κωδικοποιητής και ο αποκωδικοποιητής αποτελούνται συνήθως από ίσο αριθμό κρυφών στρωμάτων τα οποία μειώνονται σε μέγεθος όσο πλησιάζουν το χώρο προβολής, ο οποίος έχει συνήθως μέγεθος μικρότερο της εισόδου.