

REINFORCEMENT LEARNING

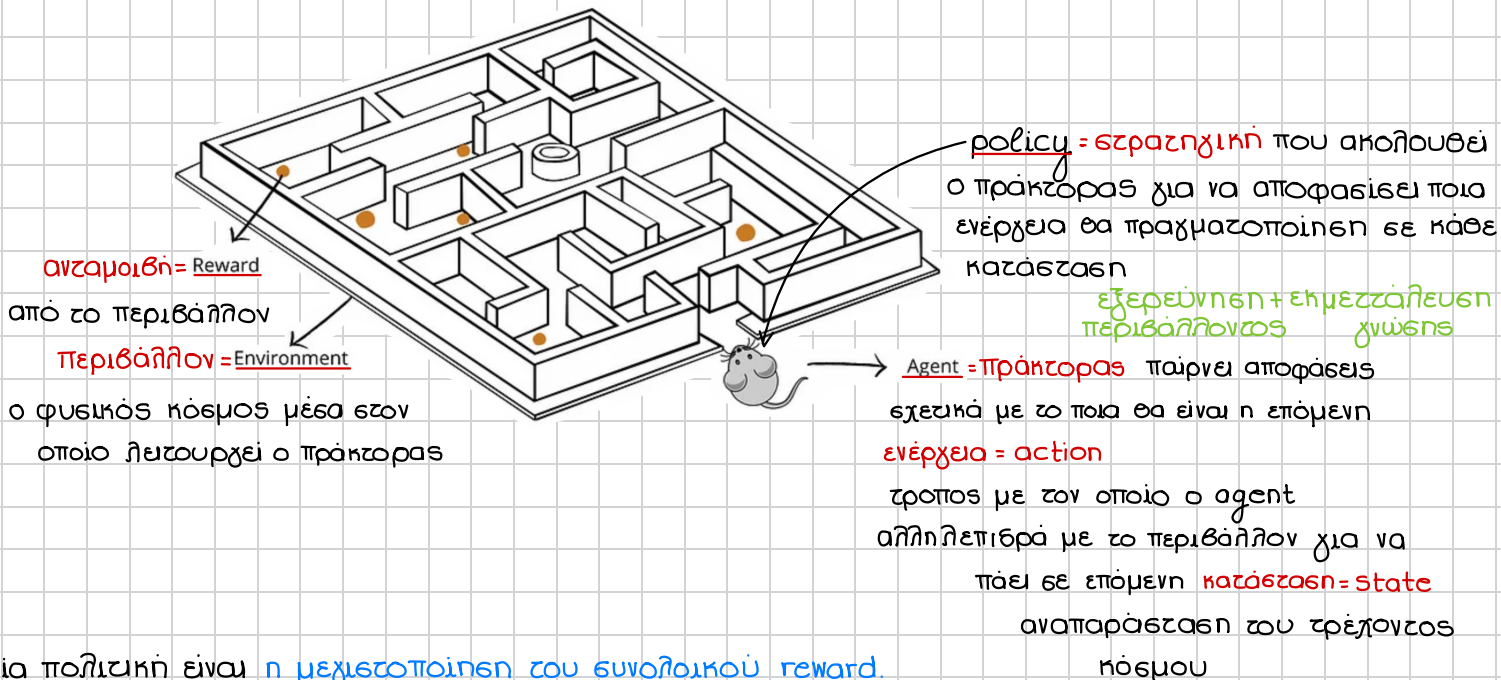
Μέχρι στιγμής έχουμε αναφερθεί σε διάφορες τεχνικές και διάφορους αλγόριθμους Machine Learning: Έχουμε δει Linear και Logistic Regression, Decision Trees, Random Forests και Ensemble Methods, Νευρωνικά Δίκτυα (CNNs και RNNs) και GANs. Οι παραπάνω τεχνικές έχουν ένα κοινό: τα δεδομένα μας εμπεριέχουν την **πραγματική έξοδο (y)** και εμείς επιθυμούμε να κατασκευάσουμε πρόγραμμα του οποίου οι **προβλεπόμενες έξοδοι (ŷ)** είναι όσο το δυνατόν πιο κοντά στις πραγματικές. (**supervised learning**)

Επίσης αναφερθήκαμε στον αλγόριθμο των k-Nearest Neighbours όπου το πρόγραμμα προσπαθεί να βρει **μία εσωτερική δομή** στα δεδομένα. (**unsupervised learning**)

Υπάρχει και ένα τρίτο είδος μάθησης:



Όροι reinforcement learning:



Μία πολιτική είναι η **μεγιστοποίηση του συνολικού reward**.

Αυτό προσπαθεί να εξασφαλίσει ο αλγόριθμος **Q-learning**.

1. Κατασκευάζουμε ένα Q-table όπου σε κάθε ζεύγος (state.action) αντιστοιχούμε μία τιμή $Q(s,a)$. **αρχικά τυχαία**.
2. Οι $Q(s,a)$ τιμές **ενημερώνονται** σύμφωνα με την **εξίσωση Bellman**: $Q(s,a) \leftarrow Q(s,a) + \alpha [r + \gamma \max_{a'} Q(s,a') - Q(s,a)]$
3. Η πολιτική καθορίζει ποια ενέργεια θα επιλεγεί σε κάθε κατάσταση. Στις περισσότερες περιπτώσεις επιλέγεται η ενέργεια που αντιστοιχεί στην μεγαλύτερη τιμή Q, αλλά μπορεί να επιλεγεί διαφορετική ενέργεια αν θέλουμε να πραγματοποιήσουμε εξερεύνηση.

