REINFORCEMENT LEARNING

Μέχρι συχμής έχουμε αναφερθεί σε διάφορες τεχνικές και διάφορους αλχορίθμους Machine Learning. Eχουμε δει Linear και Logistic Regression Decision Trees. Random Forests και Ensemble Methods Νευ ρωνικά Δίκτυα (CNNs και RNNs) και GANs. Οι παραπάνω τεχνικές έχουν ένα κοινό: τα δεδομένα μας εμπεριέχουν την προχματική έξοδο (μ) και εμείο επιθυμούμε να κατακκευά συμε πρόχραμμα του οποίου οι προ βλεπόμενες έξοδοι (ŷ) είναι όσο το δυνατόν πιο κοντά στις πραχματικές. (Supervised learning) Lπiens αναφεροπκαμε ετον αλχόρισμο των k-Nearest Neighbours όπου το πρόχραμμα προεπάθει να βρει μία ε-6ωτερική δομή ετα δεδομένα (Unsupervised learning) Υπάρχει και ενα τρίτο είδος μάθησης: HE BOED TON BODYO ETIBREUDS supervised reinforcement (EVIEXUPEVN HOONEN) unsupervised ecorning = learning= learning= learns from paoaive péemo tas tries to find patterns in unlabeled labeled data OS 34 enapodirs(nhhb νοκκρασιασι Opon reinforcement learning: -<u>policy = 6zpaznzikn</u> που ακολουθεί о практорая хиа иа апофавівей шога ενέρχεια θα πραχμαζοποίηση σε κάθε avzapol6n = Reward kazászasn EXECUTION + ENDEZZONEUEN από το περιβάλλον περιβαλλοντος XVWGDS περιδοληλον = Environment Agent = πράκτορας ποίρνει αποφάεεις o quelkos koehos heed ezon οχετικά με το ποια θα είναι η επόμενη οποίο βετουρχεί ο πράκτορας evépxea = action τροπος με τον οποίο o agent σχηνηστιορά με το περιφαμμον χια να πάει σε επόμενη κατάσταση=state αναπαρασταση του τρεχοντος Μίο ποβιτική είναι η μεχιετοποίπει του ευνοβοικού reward. κόσμου Аиго проещовы va ехабфайны о айборгороз Q-learning 1. Κατασκευάζουμε ένα Q-table οπου σε κάσε ζεύχος (state action) αντιστοιχούμε μία τιμή immediate reward Q(s.a) apzika zuzaia. 2. OI Q(s.a) tipés evnyepuvova súpopuva de znv etiemen Bellman: Q(s.a) + Q(s.a) + Q learning rate 3 Η πολιτική καθορίζει ποια ενέρχεια θα επιλεχθεί σε κάθε κατάσταση

Στις περισσότερες περιπτώσεις επιθέχεται η ενέρχεια που αντιστοιχεί στην μεχαθύτερη μήνη Ο αθθά μπο

ρεί να επιλεχθεί διαφορετική ενέρχεια αν θέλουμε να πραχματοποιήσουμε εξερεύνηση. πόμε σε state s και επιλέχουμε σετιοη ο που μεχίετο σετιοη ο που μεχίετο ποιεί την τιμή ω