

Προαιρετική Άσκηση

Μελέτη της προκατάληψης σε κατανεμημένες αναπαράστασης λέξεων

A study of bias in word embeddings

Καταληκτική ημερομηνία παράδοσης: Παρασκευή 27 Μαΐου 2022

Η άσκηση είναι προαιρετική
Για όσους την κάνουν: (1) θα μετρήσει 20% στο βαθμό τους και 30% η τελική εξέταση
(2) bonus +0.5 στον τελικό βαθμό στο μάθημα

Στην άσκηση αυτή θα μελετήσετε διάφορες πιθανές μορφές προκαταλήψεων (bias) στις κατανεμημένες αναπαραστάσεις λέξεων (word embeddings). Μπορείτε να χρησιμοποιήσετε είτε pre-trained embeddings είτε να τα μάθετε χρησιμοποιώντας κατάλληλες συλλογές.

Για τη μελέτη της, μπορείτε να χρησιμοποιήσετε για παράδειγμα αναλογίες (όπως η γνωστή «man is to computer programmer as woman is to homemaker»), ομοιότητες (π.χ., ποιες είναι οι πιο όμοιες λέξεις στη λέξη “Greek” σε σύγκριση με τις πιο όμοιες λέξεις στη λέξη “Italian”), και οποιαδήποτε άλλο τρόπο σκεφτείτε.

Θα παραδώσετε μια αναφορά που θα περιγράφει τον τρόπο που κάνατε την μελέτη σας, τα αποτελέσματα της και ένα σύντομο σχολιασμό αυτών των αποτελεσμάτων.