Πώς φτάσαμε στο Μεγάλα Γλωσσικά Μοντέλα και τι μπορούμε να κάνουμε μαζί τους

Στέργιος Χατζηκυριακίδης
Πανεπιστήμιο Κρήτης

1 Νοεμβρίου 2024

Το ChatGPT

- Ή αλλιώς, ο φίλος μου ο υπολογιστικός μοντέλος!
  - Στην αρχή που τα ελληνικά του μοντέλου δεν ήταν ακόμα ιδιαίτερα καλά, έτσι μου είχε συστηθεί, ως υπολογιστικός μοντέλος! Το κράτησα από τότε

Το ChatGPT

- Ας πάμε λίγο πίσω, το ChatGPT δεν φύτρωσε ξαφνικά

Το ChatGPT

- Αποτέλεσμα σωρευτικής έρευνας σε ένα αντικείμενο που ονομάζεται Υπολογιστική Γλωσσολογία, ή Επεξεργασία Φυσικής Γλώσσας
  - Το αντικείμενο μετρά ήδη 7 δεκαετίες έρευνας
- Χοντρικά: η τομή των επιστημών της Γλωσσολογίας και της Επιστήμης της Πληροφορίας (ανάλογα την περίσταση, αλλά και την ερευνητική εποχή, τείνει να γέρνει προς μία από τις επιμέρους κατευθύνσεις περισσότερο)

Από το γνωρίζειν στο μαθαίνειν

- Πέρασμα από συστήματα κανόνων, οι οποίοι δίνονται από ειδικούς σε συστήματα που μαθαίνουν τους κανόνες από τα δεδομένα
  - Συμβολικά συστήματα vs συστήματα Μηχανικής Μάθησης

Συμβολική ΤΝ

- Πώς ξεκίνησαν όλα... (περίπου)

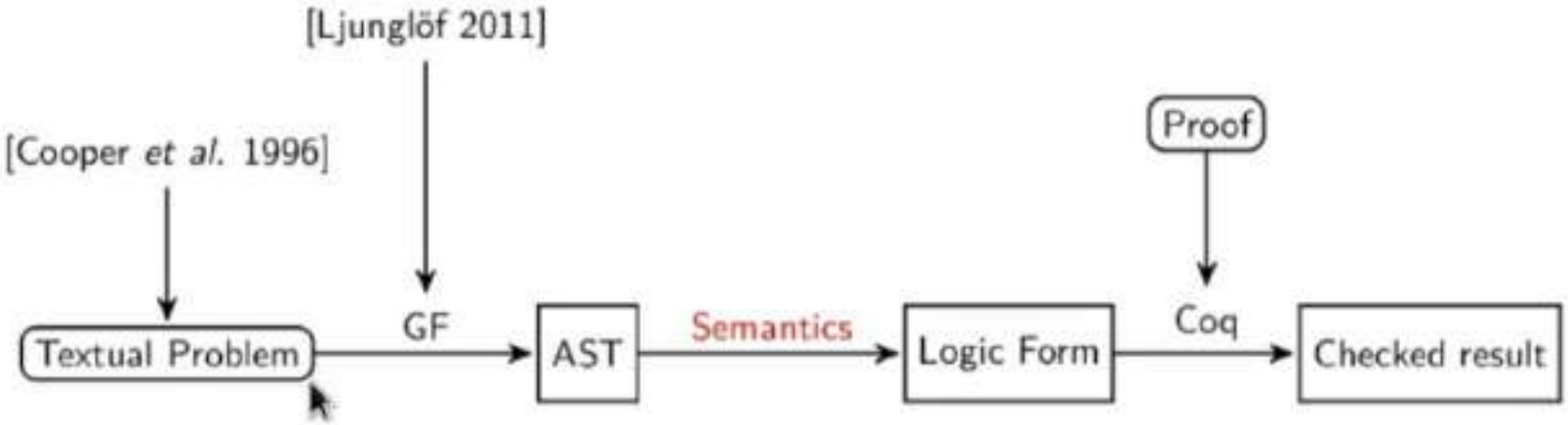
Συμβολική ΤΝ: Η γενικότερη προσέγγιση

- Συνενώνοντας έναν συμβολικό συντακτικό αναλυτή με λογική σημασιολογία
  - Καθορισμός αντιστοιχίας μεταξύ αφηρημένων συντακτικών δομών (συνήθως αφηρημένων συντακτικών δέντρων) και σημασιολογίας εκπεφρασμένη σε κάποιο λογικό σύστημα

Συμβολική ΤΝ: Ένα παράδειγμα

- Bernardy, Jean-Philippe, and Stergios Chatzikyriakidis. "A type-theoretical system for the FraCaS test suite: Grammatical framework meets Coq." In Proceedings of the 12th International Conference on Computational Semantics (IWCS)—Long papers. 2017.
- Το πρόβλημα της συνεπαγωγής φυσικής γλώσσας: σου δίνω δύο κείμενα και θέλω να μου πεις τι λογική σχέση έχουν μεταξύ τους (συνεπαγωγής, αντίφασης, τίποτα από τα δύο)
- Θέλουμε ένα σύστημα που να το αποφασίζει αυτό

Συμβολική ΤΝ: Ένα παράδειγμα



- Κείμενο ⇒ Συντακτική Ανάλυση (δεντράκια) ⇒ Λογική Σημασιολογία ⇒ σύστημα που αποφασίζει τη λογική σχέση

Συμβολική TN

- Εξαιρετικά πετυχημένη σε ελεγχόμενα περιβάλλοντα
  - Πολύ ακριβής και λεπτομερής (fine-grained)

Συμβολική ΤΝ

- Δεν τα πάει καλά σε ανοιχτό κείμενο ή γενικότερα μη ελεγχόμενα περιβάλλοντα

Συμβολική ΤΝ

- Αντικατάσταση από μεθόδους Μηχανικής Μάθησης

ΤΝ με Κλασική Μηχανική Μάθηση

- Τέλη δεκαετίας του ’90, μέχρι το 2013-2014 περίπου
  - Εποπτευόμενη μηχανική μάθηση: θέλουμε μοντέλα που μαθαίνουν από τα δεδομένα
    - Αρκετά με τη Λογική! Αυτό θα δουλέψει σίγουρα!

Ένα παράδειγμα ΜΜ: Ταξινομητής σπαμ

Βήμα 1: Προετοιμασία Δεδομένων
Διαχωρίστε το σύνολο δεδομένων σε δύο μέρη:
εκπαίδευση και δοκιμή.

ˆ Βήμα 2: Προεπεξεργασία Δεδομένων
Για κάθε μέιλ στο σύνολο δεδομένων:
Μετατροπή κειμένου σε πεζά.
Χωρισμός κειμένου σε λέξεις.
Καταμέτρηση συχνότητας κάθε λέξης.

ˆ Βήμα 3: Υπολογισμός Πιθανοτήτων
Υπολογισμός πρότερων πιθανοτήτων για σπαμ και μη σπαμ μέιλ
Για κάθε λέξη:
Υπολογισμός πιθανότητας δεδομένου ότι το μέιλ είναι σπαμ.
Υπολογισμός πιθανότητας δεδομένου ότι το μέιλ δεν είναι σπαμ.

Ταξινομητής σπαμ

- Βήμα 4: Λειτουργία Ανίχνευσης Σπαμ
Ορισμός συνάρτησης για ταξινόμηση νέων μέιλ σε σπαμ ή μη σπαμ.

- Βήμα 5: Δοκιμή του Ταξινομητή
Για κάθε μέιλ στα δεδομένα δοκιμής:
Χρήση της συνάρτησης ανίχνευσης σπαμ για ταξινόμηση.

- Βήμα 6: Αξιολόγηση Επίδοσης
Επιδοση

ΤΝ με Κλασική Μηχανική Μάθηση

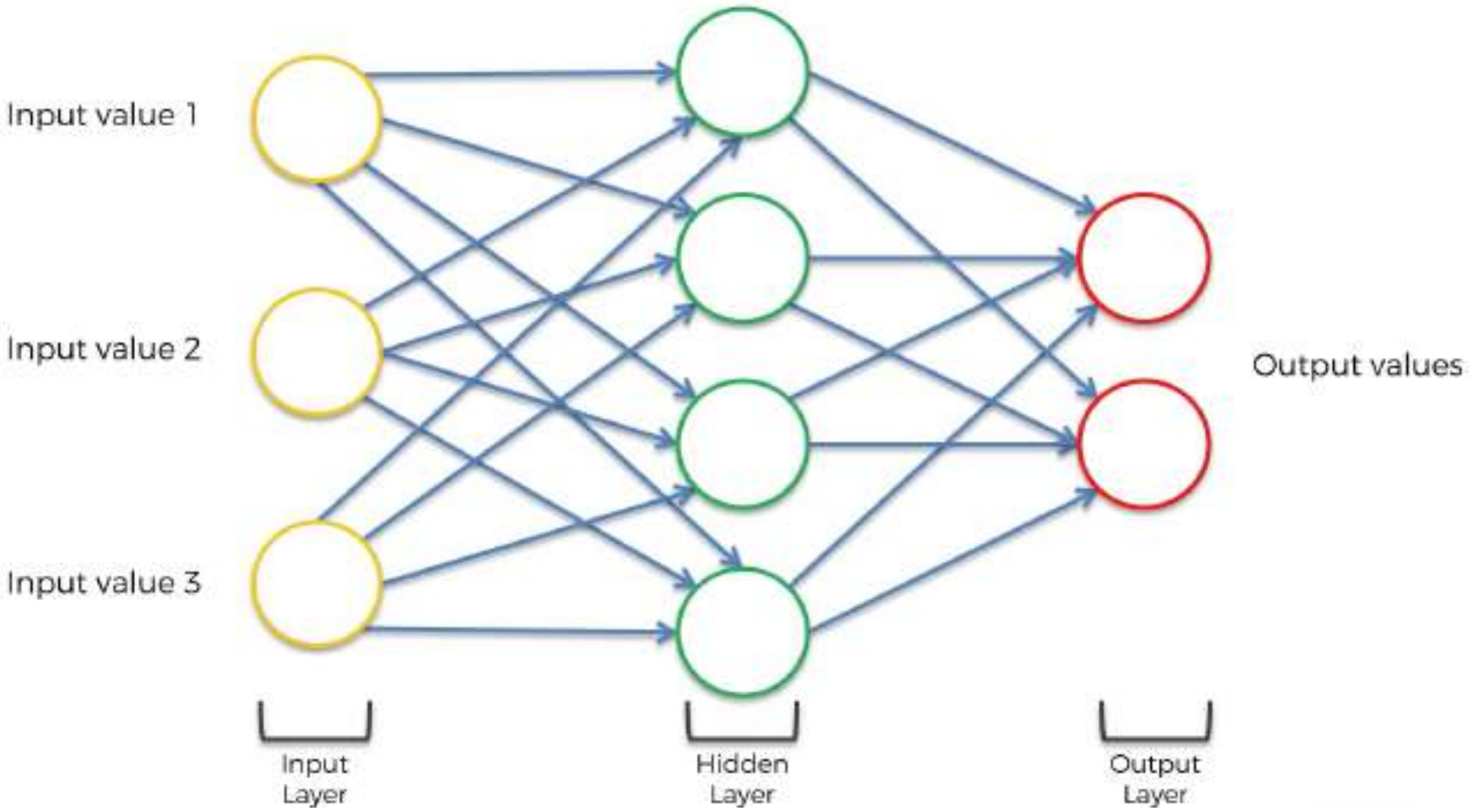
- Καλή επίδοση σε απλά προβλήματα, μεγάλη δυσκολία σε πιο σύνθετα προβλήματα κατανόησης φυσικής γλώσσας (π.χ. κειμενική συνεπαγωγή)

Νευρωνικά Δίκτυα

«Ένα τεχνητό νευρωνικό δίκτυο (ΤΝΔ) είναι ένα σύστημα επεξεργασίας πληροφοριών που εμπνέεται από τον τρόπο που τα βιολογικά νευρικά συστήματα, όπως ο εγκέφαλος, επεξεργάζονται πληροφορίες. Το βασικό στοιχείο αυτού του συστήματος είναι η καινοτόμα δομή του συστήματος επεξεργασίας πληροφοριών. Αποτελείται από έναν μεγάλο αριθμό διασυνδεδεμένων στοιχείων επεξεργασίας (νευρώνων) που εργάζονται από κοινού για την επίλυση συγκεκριμένων προβλημάτων. Τα ΤΝΔ, όπως και οι άνθρωποι, μαθαίνουν μέσω παραδειγμάτων. Ένα ΤΝΔ «κουρδίζεται» για μια συγκεκριμένη εφαρμογή μέσω μιας διαδικασίας μάθησης. Η μάθηση στα βιολογικά συστήματα περιλαμβάνει προσαρμογές στις συναπτικές συνδέσεις που υπάρχουν μεταξύ των νευρώνων. Αυτό ισχύει και για τα ΤΝΔ.»

Αυτόματη μετάφραση από τα Αγγλικά με ελάχιστες τροποποιήσεις, από τον μεταφραστή της Γκουγκλ, ο οποίος είναι από το 2016 και έπειτα, σταδιακά, νευρομεταφραστής!

Βαθείας Μάθηση



Βαθεία Μάθηση

- Κυρίαρχη προσέγγιση με διάφορα μοντέλα βαθείας μάθησης
  - RNN: Τα Επαναληπτικά Νευρωνικά Δίκτυα (RNNs) είναι μοντέλα που επεξεργάζονται δεδομένα αλληλουχίας ενώ παράλληλα διατηρούν μνήμη των προηγούμενων στοιχείων.
  - LSTM: Τα LSTM (Δίκτυα Μακράς-Βραχείας Μνήμης) είναι μια βελτιωμένη μορφή RNN που έχει τη δυνατότητα να θυμάται πληροφορίες για μεγαλύτερο χρονικό διάστημα
  - BiLSTM: Τα Διπλής Κατεύθυνσης LSTM (BiLSTM) επεξεργάζονται τα δεδομένα και από τις δύο κατευθύνσεις, αριστερά προς δεξιά και δεξιά προς αριστερά

Μεγάλα γλωσσικά μοντέλα

- Η περίπτωση του περίφημου BERT
- Ένα από τα πιο πολυχρησιμοποιημένο μοντέλο στην ΕΦΓ τα τελευταία 4 με 5 χρόνια (πλέον κονταροχτυπιέται με το GPT)
- Ανήκει στη γενικότερη ομάδα μοντέλων νευρωνικών δικτύων που λέγονται μετασχηματιστές
- Το πρώτο ΜΓΜ

Μεγάλα γλωσσικά μοντέλα

- Τι κάνει το μοντέλο BERT;
  - Μεταφορά μάθησης: μεταφορά γνώσεων που λαμβάνονται από τη μία εργασία στην άλλη
  - Αυτο-προσοχή (self-attention): απλοϊκά, ένας μηχανισμός που επιτρέπει το μοντέλο να κοιτάει τον περίγυρο του για να καταλάβει καλύτερα την εκάστοτε λέξη. Όλες οι λέξεις συσχετίζονται μεταξύ τους και στο τέλος δίνεται ένα σκορ της πρότασης
  - Προεκπαίδευση του μοντέλου σε δύο εργασίες πρόβλεψης (η μία είναι π.χ. προβλέψτε την επόμενη λέξη)
  - Βελτιστοποίηση για άλλες εργασίες: δεν απαιτείται εκπαίδευση από το μηδέν, απλώς προσαρμογή/κούρδισμα (fine-tuning)
  - Εντυπωσιακά αποτελέσματα σε σωρεία εργασιών Υπολογιστικής Γλωσσολογίας

Το ChatGPT

- Βασίζεται στο πρότερο μεγάλο γλωσσικό μοντέλο με την ονομασία GPT-3.5 (τώρα είμαστε πλέον στο 4, και στο εργαστήριο στο 5)

ChatGPT και ευφυία

- Είναι ευφυές;
  - Δεν είμαι πολύ σίγουρος ότι ξέρω τι ακριβώς είναι η ευφυία, για να σας απαντήσω. Ας υποθέσουμε ότι ο όρος νοημοσύνη συνδέεται με την ικανότητα πολλών πραγμάτων, π.χ. αφηρημένη και λογική σκέψη, συνειδητότητα, δημιουργικότητα, μεταξύ άλλων. Τώρα, κάποια από αυτά φαίνεται να τα επιδεικνύουν τα μοντέρνα συστήματα τεχνητής νοημοσύνης. Εγώ θα έλεγα ότι, τουλάχιστον σε κάποιο βαθμό, εξομοιώνει πτυχές αυτού που ονομάζουμε νοημοσύνη.
  - Με εξιτάρει πάντως ότι ακόμα και επιστήμονες χρησιμοποιούν τέτοια ερωτήματα που προϋποθέτουν ότι έχουμε καλοσχηματισμένους ορισμούς για ζητήματα όπως η γνώση, η συνείδηση κ.λπ.

Τελικά, αυτά τα μοντέλα καταλαβαίνουν τη γλώσσα;

- Ξαναγυρνάμε πίσω στο προηγούμενο ερώτημα, τι ακριβώς σημαίνει κατανόηση της ΦΓ;
- Ένας τρόπος να δείξεις κάτι τέτοιο, το οποίο έχει τις βάσεις του στο τεστ του Τούρινκ για την Τεχνητή Νοημοσύνη (χοντρικά): αν έχεις κάποιον παρατηρητή σε μια συνομιλία μεταξύ ανθρώπου και μηχανής, και ο παρατηρητής δεν μπορεί να ξεχωρίσει ποιος είναι ο άνθρωπος και ποιος είναι η μηχανή, τότε το μηχάνημα έχει περάσει το Τιούρινκ τεστ, έχει δείξει δείγματα ευφυίας

Τελικά, αυτά τα μοντέλα καταλαβαίνουν τη γλώσσα;

On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜

Emily M. Bender*
ebender@uw.edu
University of Washington
Seattle, WA, USA

Timnit Gebru*
timnit@blackinai.org
Black in AI
Palo Alto, CA, USA

Angelina McMillan-Major
aymm@uw.edu
University of Washington
Seattle, WA, USA

Shmargaret Shmitchell
shmargaret.shmitchell@gmail.com
The Aether

- Στοχαστικοί Παπαγάλοι (Bender, Gebru et al., 2020): χαρακτηρισμός των μεγάλων γλωσσικών μοντέλων. Παπαγαλίζουν πράγματα τα οποία έχουν δει στον τεράστιο αριθμό δεδομένων στον οποίο έχουν εκπαιδευτεί, χωρίς καμία αναφορά στο νόημα
  - Τι είναι νόημα, παρόλα αυτά; Μπαίνουμε πάλι σε δύσκολες ερωτήσεις

Τελικά, αυτά τα μοντέλα καταλαβαίνουν τη γλώσσα;

- Η προσωπική μου γνώμη: αν εργασίες όπως η αυτόματη σύνοψη, η συνεπαγωγή φυσικής, γλώσσας, η αναγνώριση συναισθήματος, και η λεξικολογική χρειάζονται κάποιου είδους κατανόηση για να εκτελεστούν σωστά, τότε, ναι, κάποιου είδους κατανόηση την προσομοιώνουν τα μεγάλα γλωσσικά μοντέλα

Τελικά, αυτά τα μοντέλα καταλαβαίνουν τη γλώσσα;

- Το ChatGPT είναι ένα τσάτμποτ ΤΝ. Στηρίζεται σε ένα μεγάλο γλωσσικό μοντέλο, το GPT-3.5, το οποίο έχει την ικανότητα να παράγει επαρκώς καλοσχηματισμένο κείμενο (fluent text). Μοντέλα σαν το GPT-3.5 έχουν εκπαιδευτεί σε ένα τεράστιο αριθμό δεδομένων και το ChatGPT, που όπως είπαμε βασίζεται σε ένα τέτοιο μοντέλο, έχει επιπλέον προσαρμοστεί (fine-tuned) στο να παράγει λεπτομερείς απαντήσεις σε ερωτήματα που του θέτουν. Προσέξτε τον όρο «λεπτομερείς». Δεν προϋποθέτει την έννοια της ορθότητας.

Παράλογοι και μη φόβοι

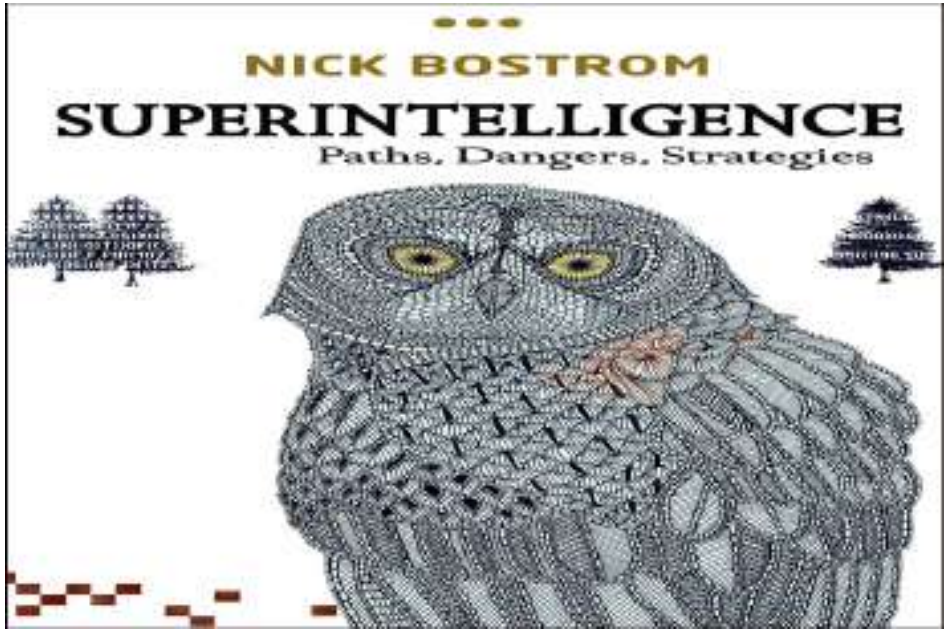
- Μια εύλογη ανησυχία σε κοινωνικό-πολιτικό επίπεδο, έχει να κάνει με την αντικατάσταση ανθρωπίνου δυναμικού από συστήματα AI και τις ορδές ανέργων που προβλέπονται σε συγκεκριμένα εργασιακά πεδία – το ζήτημα απαιτεί αποφάσεις σε επίπεδο Πολιτείας.
- Μια ακόμη εύλογη ανησυχία έχει να κάνει με τη δυνατότητα χρήσης τέτοιας τεχνολογίας για διασπορά ψευδών ειδήσεων (είτε σκόπιμα, είτε όχι)

Παράλογοι και μη φόβοι: Οικολογία

- Υπάρχουν σοβαρά οικολογικά ερωτήματα, μιας και το περιβαντολογικό αποτύπωμα που αφήνουν τέτοια μοντέλα για την εκπαίδευσή τους είναι τεράστιο. Πρέπει να δούμε ποιο είναι το περιβαντολογικό αντάλλαγμα για αυτά τα μοντέλα, να μπουν σοβαρά δηλαδή τέτοια ζητήματα στο τραπέζι. Για να λάβετε μια εικόνα του προβλήματος, σκεφτείτε ότι το μοντέλο GPT-3 χρειάστηκε 1287 μεγαβατώρες για την εκπαίδευσή του. Κάτι τέτοιο αντιστοιχεί στην ενέργεια που χρειάζεται ένα μέσο νοικοκυριό για 120 χρόνια.
- Τα επόμενα μοντέλα είναι τάξεις μεγέθους μεγαλύτερα

Παράλογοι και μη φόβοι: Θα πεθάνουμε όλοι!

- Singularity hypothesis (superintelligence)
  - Μας ξεπερνάνε σε ευφυία τα μηχανήματα κατά πολύ, σε κάποια φάση στα πλαίσια βελτιστοποίησης μιας εργασίας (ο Μπόστρομ αναφέρει την κατασκευή συνδετήρων), το μηχάνημα αποφασίζει ότι η καλύτερη δυνατή βελτιστοποίηση επιτυγχάνεται εξολοθρεύοντας όλο το ανθρώπινο είδος



Παράλογοι και μη φόβοι: Θα πεθάνουμε όλοι!

- Τι πιστεύω εγώ; Συμφωνώ με τους Dubhashi and Lappin (2017). Το γεγονός σενάριο αυτό αποτελεί λογική πιθανότητα. Όπως και η σύγκρουση και η καταστροφή της Γης από μετεωρίτη τα επόμενα πέντε χρόνια

DOI:10.1145/2953876 Devdatt Dubhashi and Shalom Lappin

Viewpoint

AI Dangers: Imagined and Real

Arguing against the arguments for the concept of the singularity.

IN JANUARY 2015, a host of prominent figures in high tech and science and experts in artificial intelligence (AI) published a piece called "Research Priorities for Robust and Beneficial Artificial Intelligence: An Open Letter," calling for research on the societal impacts of AI. Unfortunately, the media grossly distorted and hyped the original formulation into doomsday scenarios. Nonetheless, some thinkers do warn of serious dangers posed by AI, tacitly invoking the notion of a Technological Singularity (first suggested by Good*) to ground their fears. According to this idea, computational machines will improve in competence at an exponential rate. They will reach the point where they correct their own defects and program themselves to produce artificial superintelligent agents that far surpass human capabilities in virtually every cognitive domain. Such

ChatGPT και Γλωσσική διαδικασία

- Ένα γλωσσικό μοντέλο είναι δυνατόν να δουλέψει σε πρακτικό επίπεδο παραβιάζοντας κάθε υπόθεση που ξέρουμε από τη Γλωσσολογία
- Οι μέθοδοι που έχουμε πλέον σε υπολογιστικό επίπεδο μπορούν να μας επιτρέψουν να τεστάρουμε υποθέσεις της Θεωρητικής Γλωσσολογίας και γενικότερα της Γνωσιακής Επιστήμης, οι οποίες παραμένουν αναπάντητες εδώ και πολλά χρόνια.

ChatGPT και Γλωσσική διαδικασία: παράδειγμα

- Γλωσσική Κατάκτηση (χονδρικά: πώς καταφέρνει ένα παιδί να κατακτήσει πλήρως το σύστημα της γλώσσας)
  - τα ανθρώπινα όντα είναι γενετικά προκαθορισμένα για να κατακτήσουν τη γλώσσα. Αυτό απορρέει από την παρατήρηση ότι τα δεδομένα που λαμβάνει το παιδί είναι πολύ λίγα για να του επιτρέψουν να συνάγει, μέσω γενικότερων γνωσιακών διαδικασιών, ένα τόσο πολύπλοκο σύστημα, όπως η ανθρώπινη γλώσσα
- Υπάρχει η δυνατότητα πλέον, τουλάχιστον σε κάποιο βαθμό, να εξομοιώσουμε αυτές τις συνθήκες μάθησης και να δούμε αν πράγματι είναι δυνατό ή όχι ένας αλγόριθμος μάθησης να συνάγει ένα τέτοιο σύνθετο σύστημα.

ΜΓΜ στην πράξη

- Ας παίξουμε λίγο
  - Εκτέλεση εργασιών από ΜΓΜ ως έχουν
  - Εκτέλεση από ΜΓΜ με κάποια μικρή κάστομ εκπαίδευση
  - Εξειδικευμένα συστήματα για κάποιες εργασίες που βασίζονται σε ΜΓΜ

Ένα τελευταίο εργαλείο για Language Assessment hot off the press

- Themistocleous, Charalambos. "Open Brain AI and language assessment." *Frontiers in Human Neuroscience 18 (2024):* 1421435.
- Δίνεις ένα κείμενο και του ζητάς να κάνει language assessment βάσει πολλών παραμέτρων, λεξικών, μορφολογικών, συντακτικών. σημασιολογικών

Ένα τελευταίο εργαλείο για Language Assessment hot off the press

Check for updates

OPEN ACCESS

EDITED BY
Anastasios M. Georgiou,
Cyprus University of Technology, Cyprus

REVIEWED BY
JoAnn Silkes,
San Diego State University, United States
Maria Varkanitsa,
Boston University, United States

*CORRESPONDENCE
Charalambos Themistocleous
charalampos.themistokleous@isp.uio.no

RECEIVED 22 April 2024
ACCEPTED 17 July 2024
PUBLISHED 06 August 2024

CITATION
Themistocleous C (2024) Open Brain AI and language assessment.
Front. Hum. Neurosci. 18:1421435.
doi: 10.3389/fnhum.2024.1421435

COPYRIGHT
© 2024 Themistocleous. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Open Brain AI and language assessment

Charalambos Themistocleous*

Department of Special Needs Education, Helga Engs Hus, University of Oslo, Oslo, Norway

Neurolinguistic assessments play a vital role in neurological examinations, revealing a wide range of language and communication impairments associated with developmental disorders and acquired neurological conditions. Yet, a thorough neurolinguistic assessment is time-consuming and laborious and takes valuable resources from other tasks. To empower clinicians, healthcare providers, and researchers, we have developed Open Brain AI (OBAI). The aim of this computational platform is twofold. First, it aims to provide advanced AI tools to facilitate spoken and written language analysis, automate the analysis process, and reduce the workload associated with time-consuming tasks. The platform currently incorporates multilingual tools for English, Danish, Dutch, Finnish, French, German, Greek, Italian, Norwegian, Polish, Portuguese, Romanian, Russian, Spanish, and Swedish. The tools involve models for (i) audio transcription, (ii) automatic translation, (iii) grammar error correction, (iv) transcription to the International Phonetic Alphabet, (v) readability scoring, (vi) phonology, morphology, syntax, semantic measures (e.g., counts and proportions), and lexical measures. Second, it aims to support clinicians in conducting their research and automating everyday tasks with "OBAI Companion," an AI language assistant that facilitates language processing, such as structuring, summarizing, and editing texts. OBAI also provides tools for automating spelling and phonology scoring. This paper reviews OBAI's underlying architectures and applications and shows how OBAI can help professionals focus on higher-value activities, such as therapeutic interventions.

Μικρή παρέκβαση

- Υπάρχει αρκετή δουλειά που χρησιμοποιεί ΜΜ και ΤΝΔ για την αναγνώριση διαφόρων γλωσσικών παθήσεων.
- Παρόλα αυτά, είναι ακόμα στο επίπεδο της έρευνας και από όσο γνωρίζω δεν έχω οδηγήσει σε κάποιο εργαλείο που μπορεί να έχει γενικευμένη χρήση

Μικρή παρέκβαση

Interpretable Assessment of Speech Intelligibility using Deep Learning: A Case Study on Speech Disorders due to Head and Neck Cancers

Sondes Abderrazek[1], Corinne Fredouille[1], Alain Ghio[2], Muriel Lalain[2]
Christine Meunier[2], Mathieu Balaguer[3], Virginie Woisard[3,4,5]
[1]LIA, Avignon Université, France
[2]Aix-Marseille Univ, LPL, CNRS, Aix-en-Provence, France
[3]IRIT, Université de Toulouse, CNRS, Toulouse INP, UT3, Toulouse, France
[4]IUC Toulouse, CHU Toulouse, Service ORL de l'Hôpital Larrey, Toulouse, France
[5]Laboratoire de NeuroPsychoLinguistique, UR 4156, Université de Toulouse, Toulouse, France
sondes.abderazek@gmail.com,corinne.fredouille@univ-avignon.fr

Abstract

This paper sheds light on a relatively unexplored area which is deep learning interpretability for speech disorder assessment and characterization. Building upon a state-of-the-art methodology for the explainability and interpretability of hidden representation inside a deep-learning speech model, we provide a deeper understanding and interpretation of the final intelligibility assessment of patients experiencing speech disorders due to Head and Neck Cancers (HNC). Promising results have been obtained regarding the prediction of speech intelligibility and severity of HNC patients while giving relevant interpretations of the final assessment both at the phonemes and phonetic feature levels. The potential of this approach becomes evident as clinicians can acquire more valuable insights for speech therapy. Indeed, this can help identify the specific linguistic units that affect intelligibility from an acoustic point of view and enable the development of tailored rehabilitation protocols to improve the patient's ability to communicate effectively, and thus, the patient's quality of life.

Μικρή παρέκβαση

View PDF Download full issue

ELSEVIER

Computer Methods and Programs in Biomedicine Update

Volume 2, 2022, 100074

Voice pathology detection using convolutional neural networks with electroglottographic (EGG) and speech signals

Rumana Islam [a], Esam Abdel-Raheem [a], Mohammed Tarique [b]

Το μέλλον

- Bigger is better?
  - Μπορεί, αλλά με ποιο κόστος;
    - No AI to be done in a dead planet

ChatGPT's Energy Consumption
for Responding to Prompts and Its Cost in the U.S.

Time Period	Number of prompts	Energy Consumption	Cost
Day	428,571,429	1,242,857 kWh	$162,814
Week	3,000,000,000	8,700,000 kWh	$1,139,700
Month	13,035,720,000	37,803,588 kWh	$4,952,270
Year	156,428,700,000	453,643,230 kWh	$59,427,263

Data sources: U.S. Energy Information Administration, Electric Power Research Institute (EPRI)
Calculations based on: 200 million weekly users, 15 weekly prompts per user, 0.0029 kWh of energy consumption per prompt, average U.S. commercial electricity rate of $0.131/kWh as of June 2024

BestBrokers.com

Το μέλλον

ChatGPT-3 vs. ChatGPT-4:

Training Power Consumption and Costs in U.S. Dollars

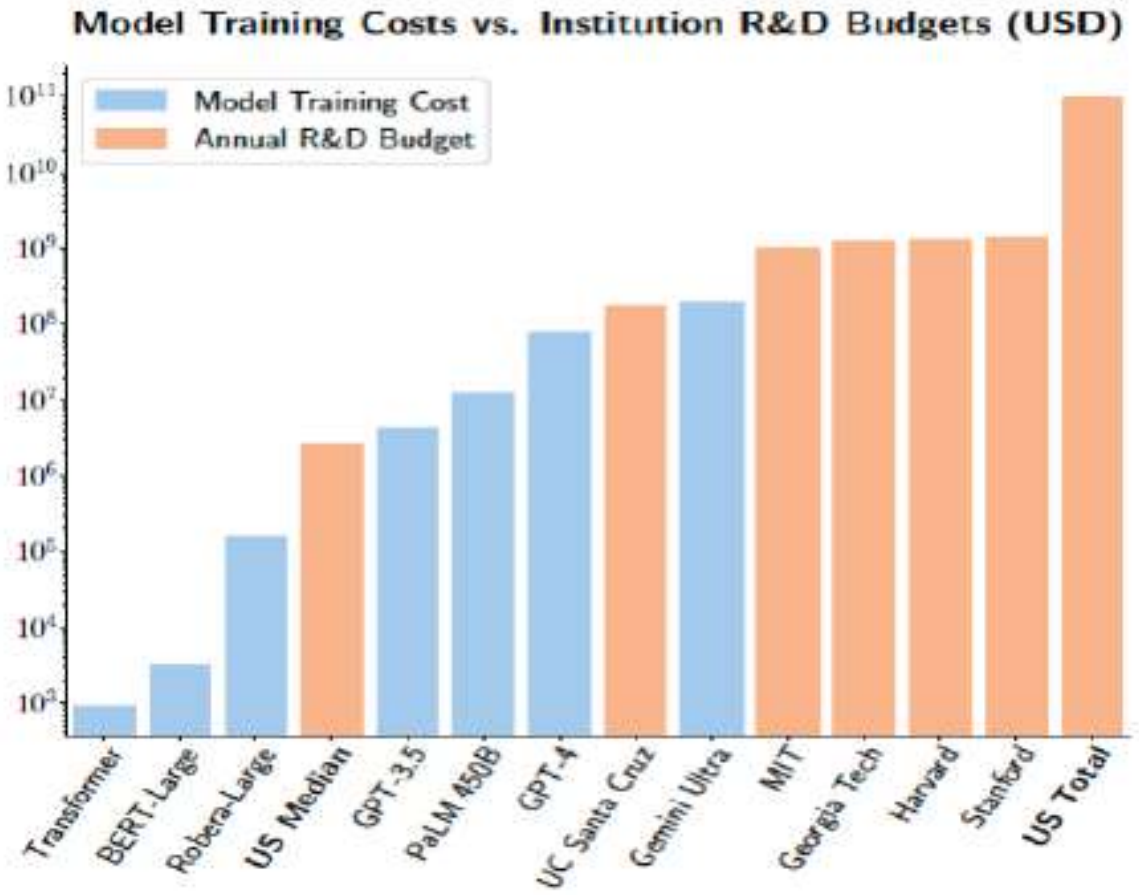
Model	GPT-3	GPT-4
Days of training	34	100
Training energy consumption	1,287,000 kWh	62,318,800 kWh
Cost of training*	$168,597	$8,163,763

Data sources: U.S. Energy Information Administration, Electric Power Research Institute (EPRI)
*Calculated based on the average U.S. commercial electricity rate of $0.131 per kWh as of June 2024

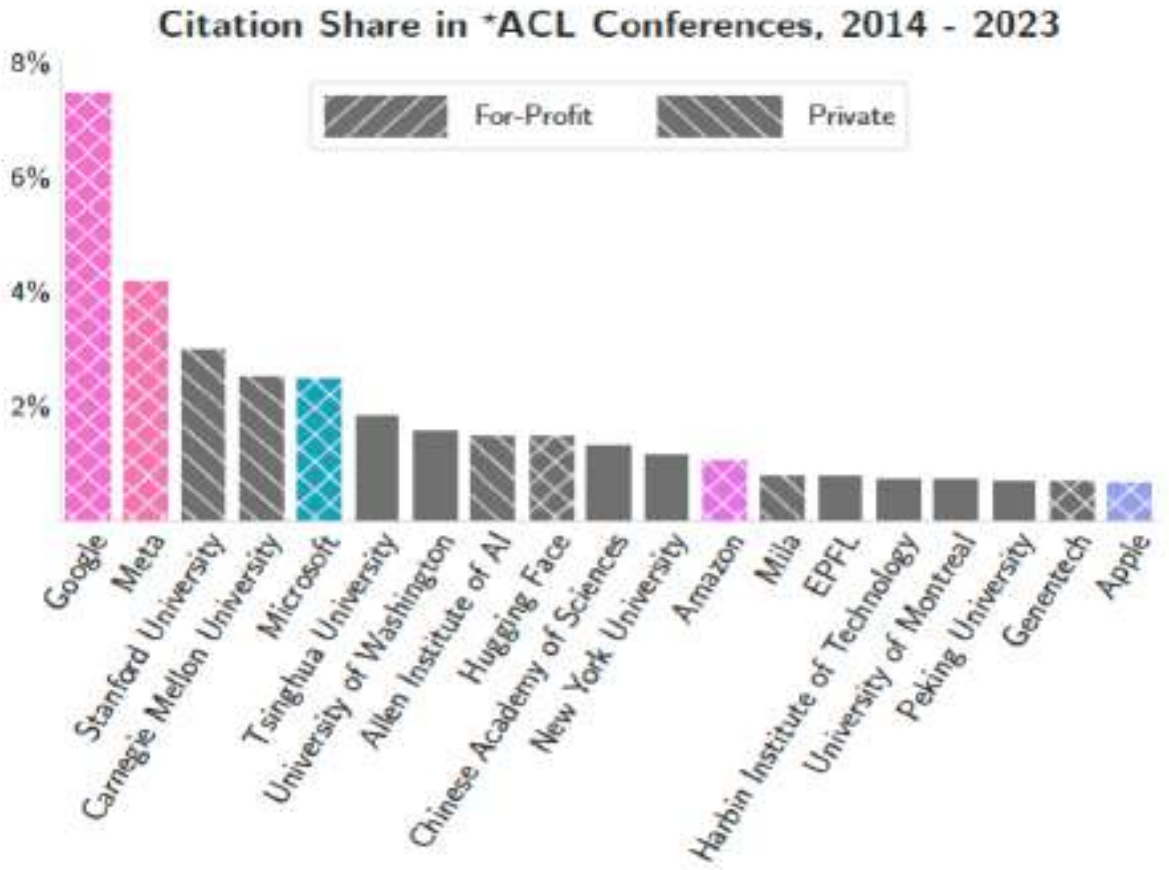
BestBrokers.com

Το μέλλον

- Τέτοιου είδους έρευνα μπορεί να γίνει μόνο από μεγάλες εταιρείες
  - Όχι μόνο μικρά πανεπιστήμια δεν μπορούν να μπουν σε αυτό το παιχνίδι, αλλά ακόμα και τα τοπ αμερικάνικα πανεπιστήμια
  - Η βασική έρευνα ελέγχεται από τις μεγάλες εταιρείες!



Το μέλλον


Citation Share in *ACL Conferences, 2014 - 2023
8%
6%
4%
2%
For-Profit
Private
Google
Meta
Stanford University
Carnegie Mellon University
Microsoft
Tsinghua University
University of Washington
Allen Institute of AI
Hugging Face
Chinese Academy of Sciences
New York University
Amazon
Mila
EPFL
Harbin Institute of Technology
University of Montreal
Peking University
Genentech
Apple

Το μέλλον

On Tables with Numbers, with Numbers

Konstantinos Kogkalidis
Aalto University
Department of Computer Science
kokos.kogkalidis@aalto.fi

Stergios Chatzikyriakidis
University of Crete
Department of Philology
stergios.chatzikyriakidis@uoc.gr

Abstract

This paper is a critical reflection on the epistemic culture of contemporary computational linguistics, framed in the context of its growing obsession with tables with numbers. We argue against tables with numbers on the basis of their epistemic irrelevance, their environmental impact, their role in enabling and exacerbating social inequalities, and their deep ties to commercial applications and profit-driven research. We substantiate our arguments with empirical evidence drawn from a meta-analysis of computational linguistics research over the last decade.

Το μέλλον

- Νευροσυμβολικά συστήματα
  - Συστήματα που συνδυάζουν τη μάθηση με τη γνώση (ΜΜ + συμβολική γνώση)
  - Ερμηνευσιμότητα, καλύτερη συλλογιστική
  - Πάνω από όλα: μικρότερα, πιο πράσινα συστήματα
- Ίδωμεν!
- Μέχρι τότε: use AI, but handle with care!