

| | |
|--|------------------------------|
| Θέμα: Ανάπτυξη και Αξιολόγηση Μοντέλων Ανάλυσης Συναισθήματος για Γλωσσικά Κείμενα Μεγάλων Δεδομένων | |
| Επιβλέπων: Ταμπακάς Βασίλης – Μ. Πιντέλας | e-mail: |
| Μέλη: | Ακαδημαϊκό Έτος: 2024 |
| <p>Στόχοι</p> <p>Η παρούσα διπλωματική εργασία εστιάζει στην ανάπτυξη και αξιολόγηση προηγμένων μοντέλων ανάλυσης συναισθήματος, ειδικά σχεδιασμένων για την επεξεργασία και ερμηνεία γλωσσικών κειμένων, με στόχο την κατανόηση και την εφαρμογή τους στον πρακτικό τομέα.</p> | |
| <p>Αντικείμενο</p> <p>Στην εποχή της ψηφιακής επικοινωνίας και των μεγάλων δεδομένων, η ανάλυση του συναισθήματος από γραπτά κείμενα αποκτά ολοένα αυξανόμενη σημασία. Καθώς η διαδικτυακή γλώσσα εξελίσσεται και γίνεται πιο πολυπλοκότερη με τη χρήση εικονιδίων, ειρωνείας και άλλων μορφών μη προφορικής επικοινωνίας, οι παραδοσιακές μέθοδοι ανάλυσης συναισθήματος γίνονται λιγότερο αποτελεσματικές. Αυτό αναδεικνύει την ανάγκη για πιο προηγμένες τεχνικές και μεθόδους στην επεξεργασία φυσικής γλώσσας (NLP) [1].</p> <p>Η συγκεκριμένη διπλωματική εργασία έχει ως αντικείμενο την ανάπτυξη και την αξιολόγηση αλγορίθμων που μπορούν να ερμηνεύσουν αποτελεσματικά το συναίσθημα από γραπτά κείμενα, όπως δημοσιεύσεις σε πλατφόρμες κοινωνικών δικτύων ή άλλες πηγές δεδομένων. Η εργασία θα επικεντρωθεί στη χρήση τεχνικών Επεξεργασίας Φυσικής Γλώσσας (NLP) [1] και προηγμένων αρχιτεκτονικών όπως οι Transformer [2] για τη δημιουργία και εκπαίδευση προ-εκπαιδευμένων μοντέλων σε γλωσσικά σύνολα δεδομένων αλλά και χρήση σύγχρονων κατανεμημένων συστημάτων μεγάλων δεδομένων (π.χ. spark / pytorch) ([3], [4]).</p> <p>[1] Qiu, X., Sun, T., Xu, Y., Shao, Y., Dai, N., & Huang, X. (2020). Pre-trained models for natural language processing: A survey. <i>Science China Technological Sciences</i>, 63(10), 1872-1897.</p> <p>[2] Lin, T., Wang, Y., Liu, X., & Qiu, X. (2022). A survey of transformers. <i>AI Open</i>.</p> <p>[3] pytorch-sentiment-analysis VS spark-nlp (https://www.libhunt.com/compare-pytorch-sentiment-analysis-vs-spark-nlp)</p> <p>[4] Next Generation Machine learning with Spark, Book, 2020</p> | |
| <p>Η εργασία περιλαμβάνει:</p> <ul style="list-style-type: none"> • Συλλογή Δεδομένων: Συλλογή και προ-επεξεργασία συνόλων δεδομένων γλωσσικού κειμένου από διάφορες πηγές (π.χ., κοινωνικά δίκτυα, δημοσιεύσεις, κριτικές προϊόντων). | |

- **Ανάλυση και Μοντελοποίηση:** Ανάπτυξη και εκπαίδευση μοντέλων NLP, με έμφαση σε προηγμένες αρχιτεκτονικές όπως οι Transformer.
- **Αξιολόγηση Μοντέλων:** Αξιολόγηση της απόδοσης των μοντέλων σε διαφορετικά σύνολα δεδομένων και σε διαφορετικές παραμέτρους.
- **Εφαρμογή και Δοκιμή:** Εφαρμογή των μοντέλων σε πραγματικά σενάρια, ανάλυση των αποτελεσμάτων και διατύπωση συμπερασμάτων.