

Διαδικτυακή εφαρμογή για την ανάλυση συναισθηματικής πληροφορίας

Νικόλαος Βρύζας^{1*}, Λάζαρος Βρύζης¹, Μαρίνα Σταματιάδου¹, Χαράλαμπος Δημούλας¹

¹ Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

*nvrzas@auth.gr (υπεύθυνος)

ΠΕΡΙΛΗΨΗ

Η παρούσα εργασία πραγματοποιείται στο πλαίσιο του έργου SHAZAAM (*Science Hoaxes to Avoid Alienation in Generation Z - A Media literacy Approach*). Το SHAZAAM αποσκοπεί στην υποστήριξη των νέων που ανήκουν στη γενιά Z για την καταπολέμηση της διάδοσης ψευδοεπιστημονικού περιεχομένου. Από τη βιβλιογραφία σε σχέση με το φαινόμενο της παραπληροφόρησης στον γραπτό λόγο, προκύπτει ότι σε μεγάλο βαθμό, το στυλ, και το συναίσθημα γραφής μπορούν να αποτελούν από μόνα τους σημαντική ένδειξη για την αναγνώριση ψευδούς περιεχομένου. Η υπόθεση είναι ότι το ίδιο ισχύει και για την οπτικοακουστική πληροφορία. Παρουσιάζεται μια διαδικτυακή εφαρμογή που παρέχει αναγνώριση συναισθήματος ως υπηρεσία, βασισμένη σε μοντέλα βαθιάς μάθησης. Η εφαρμογή έχει ως στόχο να χρησιμοποιηθεί ως εργαλείο για την ανάλυση οπτικοακουστικού περιεχομένου από τα μέσα κοινωνικής δικτύωσης.

A web application for emotion information retrieval

ABSTRACT

This work is carried out within the research project SHAZAAM (Science Hoaxes to Avoid Alienation in Generation Z - A Media literacy Approach). SHAZAAM aims to support the youth of Generation Z to combat the spread of pseudoscientific content. From the scientific literature on the phenomenon of misinformation in texts, it appears that to a large extent, the style and emotional content of writing can be an important cue in itself to identify false content. The research hypothesis is that the same applies to audiovisual content. A web application that provides sentiment recognition as a service based on deep learning models is presented. The application is intended to be used as a tool for analyzing audiovisual content from social media.

1. Εισαγωγή

Η παρούσα εργασία πραγματοποιείται στο πλαίσιο του έργου SHAZAAM (Science Hoaxes to Avoid Alienation in Generation Z - A Media literacy Approach). Το SHAZAAM αποσκοπεί στην υποστήριξη των νέων που ανήκουν στη γενιά Z για την καταπολέμηση της διάδοσης ψευδοεπιστημονικού περιεχομένου. Αυτό γίνεται με δημιουργία ομάδων εστίασης για την αναγνώριση των βασικών χαρακτηριστικών του φαινομένου, με τη δημιουργία εκπαιδευτικού υλικού, αλλά και την παροχή εργαλείων βασισμένων σε τεχνολογίες αιχμής.

Από τη βιβλιογραφία σε σχέση με το φαινόμενο της παραπληροφόρησης, προκύπτει ότι σε μεγάλο βαθμό, το στυλ, και το συναίσθημα γραφής μπορούν να αποτελούν από μόνα τους σημαντική ένδειξη για την αναγνώριση ψευδούς περιεχομένου. Κείμενα με έντονο συναισθηματικό φορτίο είναι πιο πιθανόν να έχουν ως στόχο την παραπληροφόρηση. Σχετικά πορίσματα προκύπτουν κυρίως από το πεδίο της αναγνώρισης κειμένου και ανάλυσης φυσικής γλώσσας, αλλά και από μελέτες πάνω σε πολυτροπική πληροφορία [1]-[3].

Το κίνητρο για τη διεξαγωγή της παρούσας έρευνας είναι η διερεύνηση του φαινομένου για την περίπτωση της συναισθηματικής πληροφορίας ηχητικού περιεχομένου. Η ερευνητική υπόθεση είναι ότι αφενός τα πορίσματα της βιβλιογραφίας για το φαινόμενο της παραπληροφόρησης μπορούν να επεκταθούν στην ειδική περίπτωση της ψευδοεπιστήμης και αφετέρου ότι το ζήτημα της συναισθηματικής φόρτισης ως σημαντικού κριτηρίου για την εγκυρότητα περιεχομένου έχει εφαρμογή και στην περίπτωση ηχητικού περιεχομένου. Το είδος περιεχομένου που ενδιαφέρει, ύστερα από την ανάλυση των σχετικών ομάδων εργασίας, είναι βίντεο μικρής διάρκειας που διαχέονται στα μέσα κοινωνικής δικτύωσης. Αυτό το είδος περιεχομένου είναι από τα πιο δημοφιλή στη συγκεκριμένη ηλικιακή ομάδα [3].

Στην πιο συνήθη περίπτωση, η αναγνώριση ενός τέτοιου βίντεο δεν εμπίπτει στις συνήθεις τεχνικές εντοπισμού παραποίησης περιεχομένου, καθώς τα πλάνα είναι ως επί το πλείστον αληθινά [1]. Η ανάλυση των καναλιών εικόνας, ήχου και κειμένου [1] - [4] μπορούν να συμβάλλουν στην εξαγωγή χαρακτηριστικών που να οδηγούν στην κατηγοριοποίηση σχετικού περιεχομένου. Η έλλειψη επισημειωμένων συνόλων αναφοράς αποτελεί βασικό περιορισμό στην ανάπτυξη αυτοματοποιημένων μεθόδων. Η ανάλυση συναισθήματος έχει αποδειχθεί ότι είναι μία από τις πιο κοινές τεχνικές που χρησιμοποιούν οι δημιουργοί ψευδών πληροφοριών για να απευθύνονται στο κοινό τους [5] και η γλωσσική ανάλυση προωθείται κυρίως στο [6], με την ανάλυση συναισθήματος να καλύπτει μια σημαντική πτυχή ψευδών πληροφοριών σε διάφορους τύπους περιεχομένου (ηχητικό, οπτικό, κείμενο). Ωστόσο, οι πληροφορίες σχετικά με το θέμα είναι πολύ σημαντικές για τον εντοπισμό της διάδοσης ψευδών ειδήσεων Ένας ταξινομητής μόνο για το περιεχόμενο ενός συγκεκριμένου θέματος μπορεί να οδηγήσει σε ένα μοντέλο που είναι άσχετο για μελλοντική χρήση εργαλεία ιστού για την υποστήριξη των δημοσιογράφων και του κοινού.

2. Εργαλεία αναγνώρισης ψευδοεπιστημονικών ισχυρισμών σε βίντεο κοινωνικής δικτύωσης

Αντικείμενο του προγράμματος SHAZAAM είναι η δημιουργία εργαλείων για την ενίσχυση των νέων στην αναγνώριση ψευδοεπιστημονικών ισχυρισμών σε βίντεο που διαχέονται στα μέσα κοινωνικής δικτύωσης. Τα εργαλεία που αναπτύσσονται διακρίνονται σε τρεις βασικές κατηγορίες

2.1 Εκπαιδευτικό υλικό και δράσεις μη τυπικής εκπαίδευσης

Καθώς η συγκεκριμένη μορφή περιεχομένου αναγνωρίστηκε ως επικρατέστερη στην προσέγγιση νέων σε ζητήματα ψευδοεπιστήμης, δημιουργήθηκαν στο πλαίσιο του έργου βίντεο καταπολέμησης γνωστών μύθων τα οποία παρέχονται στο κοινό μέσω της ιστοσελίδας του προγράμματος αλλά και μέσα από προγράμματα μη τυπικής εκπαίδευσης που πραγματοποιούνται σε διάφορες ακριτικές περιοχές της Ελλάδας και της Πορτογαλίας.

2.2 Εκπαιδευτικό ηλεκτρονικό παιχνίδι

Δημιουργήθηκε ένα ηλεκτρονικό παιχνίδι γνώσεων, το “Bad Gang” το οποίο απευθύνεται σε νέους της γενιάς Z και παρέχει ενημέρωση για το ζήτημα της ψευδοεπιστήμης αλλά και για γνωστούς διαδεδομένους μύθους.

2.3 Ανάλυση περιεχομένου μέσω κοινωνικής δικτύωσης

Υλοποιήθηκε διεπαφή σάρωσης περιεχομένου μέσω κοινωνικής δικτύωσης, με έμφαση στα βίντεο μικρής διάρκειας του Instagram. Η διεπαφή επιτρέπει την ανάλυση στατιστικών και τον εντοπισμό δημοφιλών θεματικών. Με αφετηρία τα επιστημονικά ευρήματα που συσχετίζουν τον εντοπισμό περιεχομένου παραπληροφόρησης με τη συναισθηματική πόλωση του περιεχομένου, η παρούσα εργασία προτείνει μια μεθοδολογία πρωτογενούς φιλτραρίσματος του υλικού αξιοποιώντας τεχνολογίες αναγνώρισης συναισθηματικής πληροφορίας σε ηχητικό περιεχόμενο.

3. Αναγνώριση συναισθηματικών μοτίβων σε οπτικοακουστικό περιεχόμενο

Το οπτικοακουστικό περιεχόμενο το οποίο ενδιαφέρει την παρούσα έρευνα μπορεί να μεταφέρει συναισθηματική πληροφορία σε διαφορετικές τροπικότητες πληροφορίας που συνυπάρχουν. Η εξαγωγή της πληροφορίας μπορεί να προσεγγιστεί σαν τα επιμέρους προβλήματα αναγνώρισης συναισθήματος στα επι

3.1 Αναγνώριση συναισθήματος μουσικής

Η μουσική είναι μία πολύ συνηθισμένη παράμετρος σε βίντεο που απευθύνονται στο συναίσθημα του θεατή καθώς είναι ένας πολύ εύκολος και αποτελεσματικός

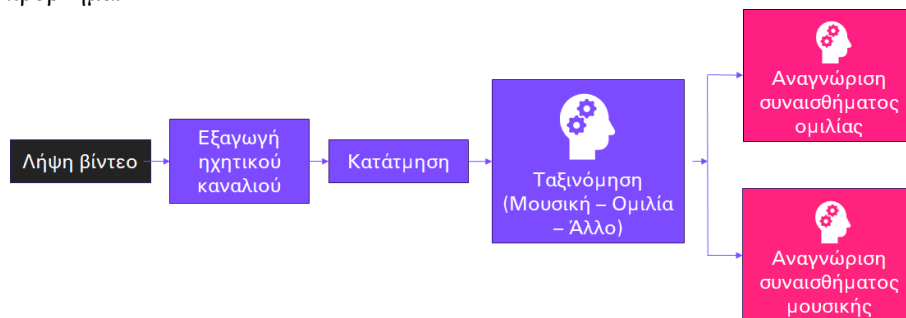
τρόπος. Για την ανάλυση της συγκεκριμένης πληροφορίας αξιοποιείται μοντέλο αρχιτεκτονικής συνελκτικού νευρωνικού δικτύου, εκπαιδευμένο στη βάση Multi-modal MIREX Emotion Dataset [8] Πρόκειται για ένα σύνολο δεδομένων που περιλαμβάνει μουσικά κομμάτια επισημειωμένα από ακροατές ως προς το προσλαμβανόμενο συναίσθημα.

3.2 Αναγνώριση συναισθήματος ομιλίας

Οποιοδήποτε ενημερωτικό βίντεο περιλαμβάνει κάποια μορφή αφήγησης. Η ερευνητική υπόθεση είναι ότι ο εντοπισμός συναισθηματικής φόρτισης σε βίντεο φερόμενα ως επιστημονικά είναι πιθανή ένδειξη ψευδοεπιστημονικού περιεχομένου. Το μοντέλο που χρησιμοποιείται είναι ένα διδιάστατο συνελκτικό νευρωνικό δίκτυο, που λαμβάνει ως δεδομένα εισόδου σπεκτρογράμματα της κλίμακας Μελ. Είναι εκπαιδευμένο στη βάση AESDD [7], που περιέχει ηχογραφήσεις ηθοποιών με συναισθηματικά φορτισμένες φράσεις σε 5 διακριτές συναισθηματικές κλάσεις.

3.3 Πολυτροπική σύντηξη πληροφορίας

Η πολυτροπική σύντηξη πληροφορίας αφορά την αξιοποίηση διαφορετικών τροπικοτήτων της οπτικοακουστικής πληροφορίας, όπως συναίσθημα μουσικής, συναίσθημα ομιλίας, συναίσθημα γλωσσικής πληροφορίας κειμένου, συναίσθημα προσώπου, συναίσθημα οπτικής πληροφορίας. Αυτή η σύντηξη μπορεί να γίνει είτε σε ανώτερο επίπεδο αξιοποιώντας μοντέλα εξειδικευμένα σε κάθε τροπικότητα ξεχωριστά, είτε δημιουργώντας ένα μοντέλο άκρη-σε-άκρη (end-to-end) που να δέχεται ως είσοδο την οπτικοακουστική πληροφορία και να δίνει μια κοινή έξοδο. Η δεύτερη περίπτωση απαιτεί την ύπαρξη ενός μεγάλου όγκου επισημειωμένων δεδομένων, κάτι που αποτελεί περιοριστικό παράγοντα για το συγκεκριμένο πρόβλημα.



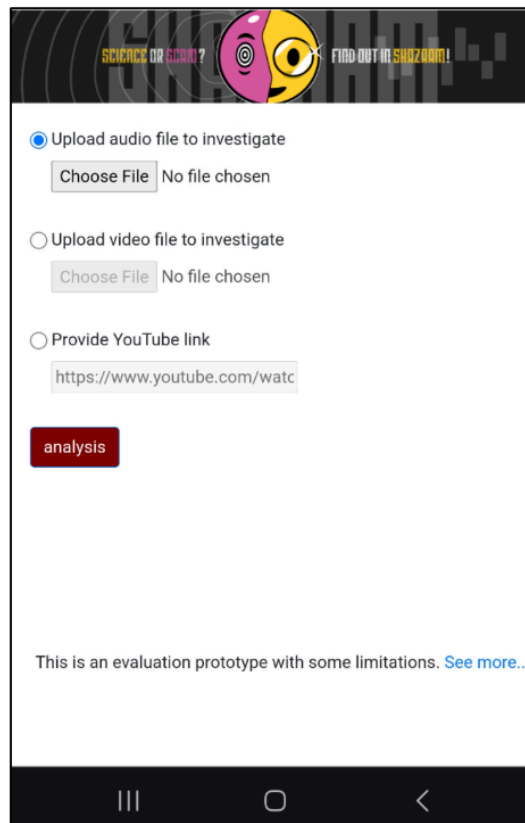
Σχήμα 4.1 Η βασική λειτουργικότητα του εργαλείου αναγνώρισης συναισθηματικής πληροφορίας

4. Διεπαφή αναγνώρισης συναισθηματικής πληροφορίας

Για την υλοποίηση της διεπαφής δημιουργήθηκε μια εικονική μηχανή η οποία λειτουργεί σαν εξυπηρετητής. Η λειτουργικότητα υλοποιήθηκε με τη βιβλιοθήκη flask σε γλώσσα python, υλοποιώντας μια αρχιτεκτονική δικτυακών υπηρεσιών. Παράλληλα με το γραφικό περιβάλλον και τη λειτουργικότητα διάδρασης με τον

χρήστη, οι δικτυακές υπηρεσίες περιλαμβάνουν την ταξινόμηση ηχητικού περιεχομένου χρησιμοποιώντας τα μοντέλα αναγνώρισης συναισθήματος μουσικής και ομιλίας που αναφέρθηκαν, καθώς και ένα στάδιο ταξινόμησης της ηχητικής πληροφορίας σε μουσική, ομιλία ή άλλους ήχους [9]. Για κάθε χρονικό παράθυρο, το ηχητικό περιεχόμενο ταξινομείται ως σήμα ομιλίας, μουσικής, ή άλλος ήχος και καλείται το αντίστοιχο μοντέλο συναισθηματικής ταξινόμησης (Σχήμα 4.1).

Το περιβάλλον αλληλεπίδρασης με τον χρήστη φαίνεται στο Σχήμα 4.2. Ο χρήστης δίνει ως είσοδο ένα αρχείο ήχου, ένα αρχείο βίντεο ή έναν σύνδεσμο στο YouTube. Στην πρώτη περίπτωση, το αρχείο ήχου καταταμείται και στέλνεται ως είσοδος στις υπηρεσίες αναγνώρισης συναισθήματος. Στην περίπτωση του βίντεο, απομονώνεται περιλαμβάνεται ένα στάδιο απομόνωσης της ηχητικής πληροφορίας, ενώ στην περίπτωση του συνδέσμου προβλέπεται ένα επιπλέον στάδιο λήψης του αρχείου.



Σχήμα 4.2 ένα στιγμιότυπο της εφαρμογής SHAZAAM

5. Συμπεράσματα και μελλοντικές επεκτάσεις

Παρουσιάστηκε ένα σύστημα αναγνώρισης συναισθηματικής πληροφορίας ηχητικού περιεχομένου. Ο στόχος είναι το σύστημα να χρησιμοποιηθεί ως ένα εργαλείο εντοπισμού ψευδοεπιστημονικού οπτικοακουστικού περιεχομένου. Η προτεινόμενη λειτουργικότητα παρέχεται και ως γραφική διεπαφή για τη χρήση χωρίς τεχνικές γνώσεις.

Οι μελλοντικές ερευνητικές κατευθύνσεις αφορούν την αξιολόγηση της μεθόδου σε σύνολα πραγματικών δεδομένων, η αξιολόγηση του εργαλείου από το κοινό απεύθυνσης και η ενσωμάτωσή του σε ένα πολυτροπικό σύστημα αναγνώρισης συναισθήματος.

6. Ευχαριστίες



Η παρούσα έρευνα υποστηρίζεται από το European Media and Information Fund. Η ευθύνη του περιεχομένου βαρύνει αποκλειστικά τους συγγραφείς και δεν αντικατοπτρίζει απαραίτητα τις θέσεις του EMIF, του Calouste Gulbenkian και του European University Institute.

7. Αναφορές

- [1] Li, Xiaojun, et al. "A CNN-based misleading video detection model." *Scientific Reports* **12.1** pp 6092 (2022).
- [2] Papadopoulou, O., et al. "A corpus of debunked and verified user-generated videos." *Online information review* **43.1**: 72-88.(2019)
- [3] Qi, P., et al. "Fakesv: A multimodal benchmark with rich social context for fake news detection on short video platforms." *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 37. No. 12. (2023).
- [4] Papadamou, K., Zannettou, S., Blackburn, J., De Cristofaro, E., Stringhini, G., & Sirivianos, M.. "It is just a flu": assessing the effect of watch history on YouTube's pseudoscientific video recommendations. In *Proceedings of the international AAAI conference on web and social media* (Vol. 16, pp. 723-734) (2022)
- [5] Alonso, M. A., Vilares, D., Gómez- Rodríguez, C., & Vilares, J.. "Sentiment analysis for fake news detection." *Electronics*, **10(11)**, 1348 (2021).
- [6] Mishra, S., Shukla, P., & Agarwal, R.. "Analyzing machine learning enabled fake news detection techniques for diversified datasets." *Wireless Communications and Mobile Computing* 2022, **1-18** (2022).
- [7] Vryzas, N., Kotsakis, R., Liatsou, A., Dimoulas, C. A., & Kalliris, G. (2018). Speech emotion recognition for performance interaction. *Journal of the Audio Engineering Society*, 66(6), 457-467.
- [8] Panda R., Malheiro R., Rocha B., Oliveira A. & Paiva R. P. "Multi-Modal Music Emotion Recognition: A New Dataset, Methodology and Comparative Analysis". *10th International Symposium on Computer Music Multidisciplinary Research -- CMMR'2013*, Marseille, France (2013)

[9] Vrysis, L., Tsipas, N., Thoidis, I., & Dimoulas, C.. “1D/2D deep CNNs vs. temporal feature integration for general audio classification.” *Journal of the Audio Engineering Society*, **68(1/2)**, 66-77. (2020)

[10] Vryzas, N., Vrysis, L., Kostarella, I., Dimoulas, C. A., “Investigating audio emotional patterns in pseudoscience” in *Proceedings of the 156th Audio Engineering Society Convention* (2024)