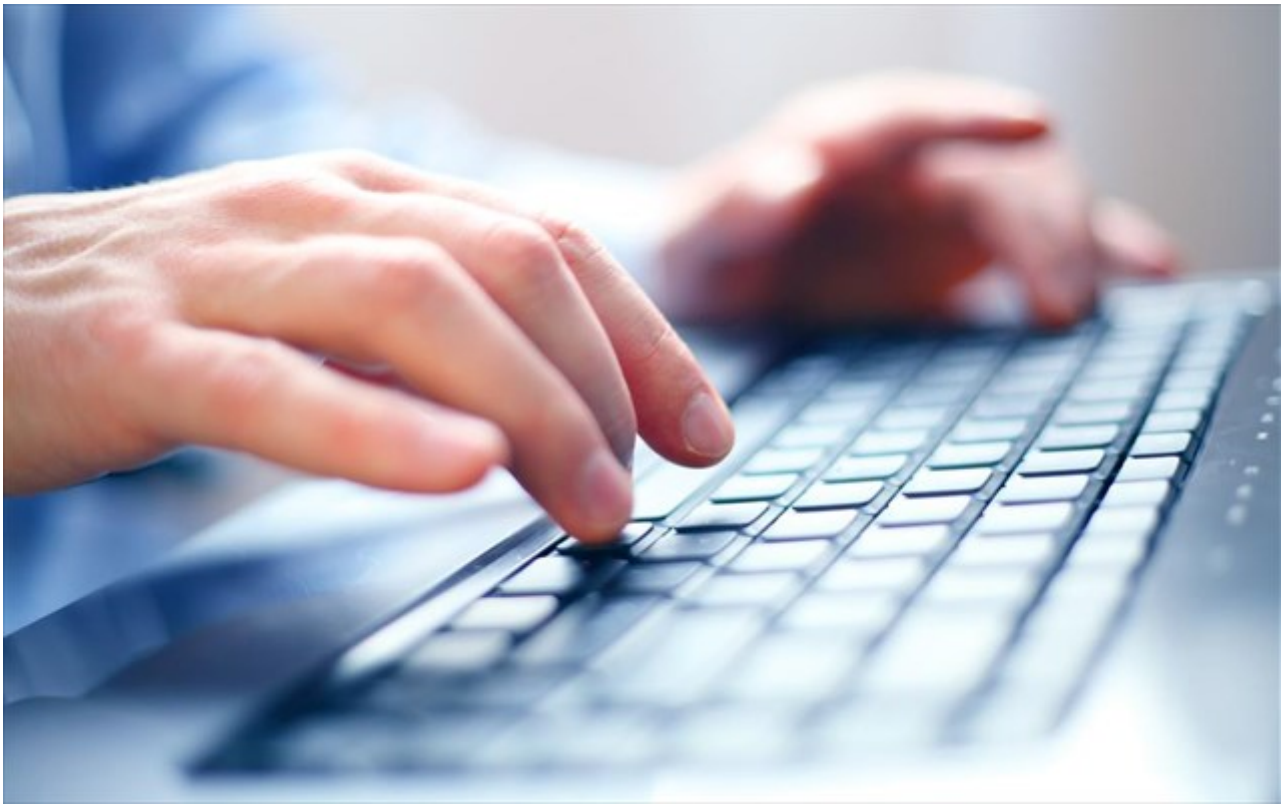
29 Απριλίου 2017

# «Κούρσα» για την ανάπτυξη λογισμικού που εντοπίζει τα βίντεο με ανεπιθύμητο υλικό στο Ίντερνετ

/ Επιστήμες, Τέχνες & Πολιτισμός





*SHUTTERSTOCK*

*Με το θέμα ασχολούνται ήδη κολοσσοί του χώρου, όπως το Facebook και η Google (που αντιμετωπίζει παρόμοια προβλήματα στο YouTube. Οι περισσότεροι επικεντρώνονται στο deep learning, που χρησιμοποιεί υπολογιστικά «νευρωνικά δίκτυα».*

Εταιρείες από όλο τον κόσμο επιδίδονται σε μια «κούρσα» για την ανάπτυξη και βελτίωση συστημάτων τεχνητής νοημοσύνης που θα εντοπίζουν και θα μπλοκάρουν αυτόματα τα βίντεο που περιέχουν φόνους και γενικότερα βίαιο περιεχόμενο, πριν αυτά εξαπλωθούν μέσω social media. Ωστόσο, όπως σημειώνει το Reuters, μέχρι τώρα καμία δεν έχει καταφέρει να λύσει απόλυτα το πρόβλημα.

Το πιο πρόσφατο παράδειγμα είναι αυτό ενός άνδρα στην Ταϊλάνδη που κατέγραψε τον εαυτό του να σκοτώνει την 11 μηνών κόρη του σε ζωντανό βίντοε στο Facebook. Το συγκεκριμένο συμβάν- του οποίου έχουν προηγηθεί αρκετά άλλα- θέτει ερωτήματα σχετικά με το πώς λειτουργεί το σύστημα αναφοράς του Facebook και το πώς θα γίνει δυνατόν να υποδεικνύεται ταχύτερα τέτοιου είδους υλικό.

Με το θέμα ασχολούνται ήδη κολοσσοί του χώρου, όπως το Facebook και η Google (που αντιμετωπίζει παρόμοια προβλήματα στο YouTube. Οι περισσότεροι επικεντρώνονται στο deep learning, που χρησιμοποιεί υπολογιστικά «νευρωνικά δίκτυα». Κατά τον Ματ Ζάιλερ, ιδρυτή και διευθύνοντα σύμβουλο της Clarifai (Νέα Υόρκη), ο τομέας αυτός έχει «απογειωθεί» μόλις τα τελευταία χρόνια. Η μέθοδος αυτή περιλαμβάνει την παροχή εικόνων στα «στρώματα νευρώνων» του υπολογιστή, ο οποίος στη συνέχεια «μαθαίνει» να αναγνωρίζει σκηνές βίας, καθώς «εκπαιδεύεται» να καταλαβαίνει βίαιες κινήσεις, αίμα κλπ, όπως αναφέρει ο

Αμπχιτζίτ Σάνμπχαγκ, διευθύνων σύμβουλος της Graymatics στη Σιγκαπούρη. Επίσης, σύμφωνα με τον Ζάιλερ, οι αλγόριθμοι της εταιρείας μπορούν επίσης να αναγνωρίζουν αντικείμενα που θα μπορούσαν να είναι «προάγγελοι» βίας, πχ μαχαίρια, όπλα κ.α.

Ωστόσο, υπάρχουν όρια: Για αρχή, το λογισμικό είναι τόσο αποτελεσματικό όσο τα παραδείγματα πάνω στα οποία εκπαιδεύεται. Εάν κάποιος πάει να κρεμάσει ένα παιδί από ένα κτίριο, αυτό δεν σημαίνει απαραίτητα ότι το πρόγραμμα έχει εκπαιδευτεί ήδη να αναγνωρίζει κάτι τέτοιο. «Καθώς οι άνθρωποι γίνονται πιο εφευρετικοί πάνω σε σε τέτοιες απεχθείς δραστηριότητες, το σύστημα πρέπει να εκπαιδευτεί σε αυτό» λέει ο Σανμπχάγκ, η εταιρεία του οποίου φιλτράρει βίντεο και εικόνες για πελάτες στον χώρο των social media στην Ασία και αλλού. Άλλο ένα πρόβλημα είναι ότι η βία μπορεί να μην είναι «αντικειμενική», πχ περιπτώσεις όπου ασκείται ψυχολογική βία, ενώ υπάρχουν και περιπτώσεις όπου μπορεί να υπάρχει αμφιλεγόμενο υλικό που δεν είναι από μόνο του «βίαιο», πχ μια σημαία του Ισλαμικού Κράτους.

Ένας άλλος περιορισμός είναι πως, όσο και αν βοηθούν οι αυτοματισμοί, απαιτούνται πάντα άνθρωποι στο «κύκλωμα», για να επιβεβαιώνουν αν το υλικό που έχει υποδειχθεί ως προσβλητικό, εξαιρετικά βίαιο ή επικίνδυνο, είναι όντως, όπως υποδεικνύει ο Μίκα Ραουτιάινεν, ιδρυτής και διευθύνων σύμβουλος της φινλανδικής Valossa, η οποία εντοπίζει ανεπιθύμητο υλικό για ΜΜΕ και διαφημιστικές εταιρείες.

Ωστόσο, ανεπιθύμητη δεν είναι μόνο η βία: Ο Μάικλ Πόγκρενμπιακ, διευθύνων σύμβουλος της Kuznech, υποστηρίζει πως η εταιρεία του έχει αναβαθμίσει το «οπλοστάσιό» της, με αλγορίθμους εντοπισμού πορνογραφικών εικόνων (που περιλαμβάνουν κυρίως εντοπισμό έκθεσης δέρματος και κίνησης), με προγράμματα που εντοπίζουν τα logos από στούντιο παραγωγής και προειδοποιητικά κείμενα. Αντίστοιχες μεθόδους χρησιμοποιεί και το Facebook, που χρησιμοποιεί παρόμοιες τεχνικές για να βρίσκει γυμνό, βία και άλλα θέματα που παραβιάζουν την πολιτική του.

**Πηγή:** naftemporiki.gr