Co-funded by the European Union

STEAME TEACHER ACADEMY

Χρηματοδοτείται από την Ευρωπαϊκή Ένωση. Οι απόψεις και οι γνώμες που εκφράζονται είναι αποκλειστικά του/των συγγραφέα/ων και δεν αντανακλούν κατ' ανάγκη τις απόψεις και τις γνώμες της Ευρωπαϊκής Ένωσης ή του Ευρωπαϊκού Εκτελεστικού Οργανισμού Εκπαίδευσης και Πολιτισμού (EACEA). Ούτε η Ευρωπαϊκή Ένωση ούτε ο EACEA μπορούν να θεωρηθούν υπεύθυνοι γι' αυτές.

STEAME ACADEMY
ΣΧΕΔΙΟ ΜΑΘΗΣΗΣ ΚΑΙ ΔΗΜΙΟΥΡΓΙΚΟΤΗΤΑΣ ΔΙΕΥΚΟΛΥΝΣΗΣ ΤΗΣ ΔΙΔΑΣΚΑΛΙΑΣ (ΣΧΕΔΙΟ ΜΑΘΗΣΗΣ ΚΑΙ ΔΗΜΙΟΥΡΓΙΚΟΤΗΤΑΣ) - ΕΠΙΠΕΔΟ 1 ΦΟΙΤΗΤΩΝ
Εξόρυξη κειμένου: είναι αυτά τα έγγραφα τα ίδια;

S ***T*** ***Eng*** ***A*** ***M*** ***Ent***

1. Επισκόπηση

Τίτλος	Εξόρυξη κειμένου: είναι αυτά τα έγγραφα τα ίδια;
Ερώτηση ή θέμα αφόρμησης	Πώς μπορούν οι μηχανές αναζήτησης να βρουν αποτελέσματα για μια αναζήτηση χρήστη με βάση λέξεις-κλειδιά; Πώς αναγνωρίζουν οι υπολογιστές έγγραφα κειμένου που εστιάζουν στα ίδια θέματα; Πώς μοντελοποιούν οι αλγόριθμοι υπολογιστών μη δομημένα δεδομένα για ψηφιακή επεξεργασία;
Ηλικίες, βαθμίδες, …	16 έως 18 ετών Α' – Γ' Λυκείου
Διάρκεια, χρονοδιάγραμμα, δραστηριότητες	132 ώρες 17 δραστηριότητες
Ευθυγράμμιση με το πρόγραμμα σπουδών	Εξόρυξη δεδομένων, μηχανική μάθηση, μοντελοποίηση μη δομημένων δεδομένων, προγραμματισμός υπολογιστών
Συνεισφέροντες, Συνεργάτες	
Περίληψη - Σύνοψη	Οι μαθητές εισάγονται στην Εξόρυξη Δεδομένων και τη Μηχανική Μάθηση εστιάζοντας στα βασικά θέματα της ψηφιακής επεξεργασίας κειμένου. Η ομοιότητα του κειμένου διερευνάται δείχνοντας τα μαθηματικά του θεμέλια, τη διασταύρωση των συνόλων και το συνημίτονο μεταξύ δύο διανυσμάτων. Οι μαθητές εργάζονται σε ομάδες για να εφαρμόσουν ένα απλό εργαλείο για τη μέτρηση της ομοιότητας μεταξύ δύο εγγράφων κειμένου. Στις τελευταίες συνεδρίες οι μαθητές καλούνται για έναν διαγωνισμό για τον προσδιορισμό της καλύτερης εφαρμογής. Κατά τη διάρκεια όλων των συνεδριών, οι μαθητές εισάγονται στις βασικές μεθόδους προ-επεξεργασίας κειμένου, όπως οι stop-words (λέξεις που είναι πολύ συχνές στο κείμενο αλλά δεν έχουν μεγάλη σημασία για την ανάλυση του περιεχομένου) και η αποκοπή καταλήξεων. Το τελευταίο μάθημα ολοκληρώνεται βάζοντας τους μαθητές να συζητήσουν και να εντοπίσουν τις ομοιότητες μεταξύ της υλοποίησής τους και μιας μηχανής αναζήτησης και, από εκεί, να σχεδιάσουν μια μηχανή αναζήτησης χρησιμοποιώντας την προηγούμενη εφαρμογή τους για την ομοιότητα

Αναφορές, παραθέσεις	κειμένου.

2. Πλαίσιο STEAME ACADEMY*

Συνεργασία των εκπαιδευτικών	Εκπαιδευτικός 1 (Φυσικές επιστήμες) • Μηχανική μάθηση και εξόρυξη δεδομένων: επισκόπηση του πεδίου, γενική αρχιτεκτονική, κοινές εφαρμογές στην καθημερινή ζωή, κοινά ζητήματα • Μοντελοποίηση δεδομένων για ψηφιακή επεξεργασία: αναπαράσταση δομημένων δεδομένων για μηχανική μάθηση και εξόρυξη δεδομένων, μοντελοποίηση μη δομημένων δεδομένων για μηχανική μάθηση και εξόρυξη δεδομένων • Εξόρυξη κειμένου: επισκόπηση του πεδίου, βασικές έννοιες, μοντελοποίηση (μοντέλο Boolean, μοντέλο TF, μοντέλο TFxIDF), ομοιότητα εγγράφων κειμένου, προεπεξεργασία, κύριες εργασίες και εφαρμογές Εκπαιδευτικός 2 (Μηχανικός) • Προγραμματισμός Python για εξόρυξη κειμένου • Διανυσματική άλγεβρα στο Excel Εκπαιδευτικός 3 (Μαθηματικά) • Συνάρτηση συνημιτόνου, διανυσματική άλγεβρα • Τελεστές ρύθμισης, διασταύρωση Ο Εκπαιδευτικός 1 συνεργάζεται με τον Εκπαιδευτικό και τον Εκπαιδευτικό 3 για να: - προσδιορίσουν τις βιβλιοθήκες Python που θα χρησιμοποιηθούν για εξόρυξη κειμένου - προσδιορίσουν τις συναρτήσεις του Excel που θα χρησιμοποιηθούν για τη διανυσματική άλγεβρα - δημιουργήσουν τις ασκήσεις και την πρόκληση Ο εκπαιδευτικός 1 συνεργάζεται με τον εκπαιδευτικό 2 για να: - συγκεντρώσουν τα απαραίτητα σώματα κειμένων για τις πρακτικές δραστηριότητες - επισημειώσουν τα απαραίτητα σώματα κειμένων για κάθε άσκηση Ο εκπαιδευτικός 1 συνεργάζεται με τον εκπαιδευτικό 3 για να: - κάνουν μία εισαγωγή στους τελεστές συνόλου σε περιβάλλον επεξεργασίας κειμένου - κάνουν μια εισαγωγή της διανυσματικής άλγεβρας σε περιβάλλον

	επεξεργασίας κειμένου.
Οργάνωση STEAME in Life (SiL)	Η τελευταία συνεδρία ολοκληρώνεται βάζοντας τους μαθητές να συζητήσουν και να εντοπίσουν τις ομοιότητες μεταξύ της υλοποίησής τους και μιας μηχανής αναζήτησης, όπως η Google, και, από εκεί, να σχεδιάσουν μια μηχανή αναζήτησης χρησιμοποιώντας την προηγούμενη εφαρμογή τους για την ομοιότητα κειμένου.
Διαμόρφωση Σχεδίου Δράσης	**Προπαρασκευαστική φάση** 1. Έρευνα, εξόρυξη δεδομένων και μηχανική μάθηση, παραδοσιακές και σύγχρονες εφαρμογές, σύνδεση με μηχανές αναζήτησης και παραγωγική τεχνητή νοημοσύνη αναφέρονται σε περιπτώσεις δομημένων και μη δομημένων δεδομένων, ανασκόπηση των κύριων προκλήσεων της εξόρυξης δεδομένων και της μηχανικής μάθησης (μοντελοποίηση δεδομένων, πολυδιάστατη, υπερπροσαρμογή, ελλιπή δεδομένα, όγκος δεδομένων, μαζικά δεδομένα, εξήγηση έναντι πρόβλεψης, ...) 2. Συλλογή και επισημείωση σωμάτων κειμένων για τις ασκήσεις 3. Ρυθμίστε το περιβάλλον προγραμματισμού Python (docker, αποθετήριο στο Github για κλωνοποίηση, άλλα) **Δομή εργαστηρίου** 1. Εισαγωγή 1.1. Επισκόπηση Εξόρυξης Δεδομένων και Μηχανικής Μάθησης: ιστορική προοπτική, εργασίες/προβλήματα, εφαρμογές, μοντελοποίηση μη δομημένων δεδομένων όπως εικόνες και κείμενο. 2. Μοντελοποίηση κειμένου 2.1. Συγκεντρώστε τα σώματα κειμένων που θα χρησιμοποιηθούν για τις ασκήσεις και την εργασία του έργου (πρέπει να έχουν μικρά έγγραφα με πολύ μειωμένο λεξικό, συμπεριλάβετε έγγραφα όπου το TF μπορεί να κάνει τη διαφορά σε σύγκριση με το μοντέλο Boolean. 2.2. Μοντέλα για την αναπαράσταση κειμένου, παραδείγματα: Μοντέλο Boolean, TF, TFxIDF, άλλο. Ανατρέξτε στο Bag of Words. 2.3. Ορισμός τελεστών (ένωση, διασταύρωση κ.λπ.). 2.4. Εφαρμόστε διάφορες εκδόσεις ενός εργαλείου στο Excel για να υπολογίσετε την ομοιότητα μεταξύ δύο εγγράφων κειμένου με βάση τελεστές συνόλου ή / και διανυσματική άλγεβρα (εσωτερικό γινόμενο, συνημίτονο, τομή και ένωση συνόλων, ...). 2.5. Εισαγωγή διανυσματικής άλγεβρας: εσωτερικό γινόμενο και συνημίτονο. 2.6. Δείξτε τον αντίκτυπο/συνάφεια της συχνότητας των όρων σε σύγκριση με το Boolean. Καταστήστε σαφές ότι το συνημίτονο λειτουργεί καλά. 2.7. Προκλήσεις ομοιότητας που βασίζονται μόνο στο TF (όροι που υπάρχουν σε όλα τα έγγραφα δεν έχουν εξουσία διάκρισης), εναλλακτικές λύσεις. 2.8. Παρουσιάστε τα IDF και TFxIDF συζητήστε, σχεδιάστε και υλοποιήστε στο Excel ένα μέτρο ομοιότητας βασισμένο στο μοντέλο TFxIDF που θα

χρησιμοποιηθεί για έγγραφα με μειωμένο λεξικό.

3. Εφαρμογή
   3.1. Ερευνήστε την R, την Python ή άλλες βιβλιοθήκες για εξόρυξη κειμένου. Προετοιμάστε ένα αποθετήριο και έναν οδηγό εγκατάστασης για τους μαθητές ώστε να εγκαταστήσουν αυτό το περιβάλλον ανάπτυξης.
   3.2. Υλοποιήστε μια συνάρτηση για να υπολογίσετε την ομοιότητα μεταξύ δύο εγγράφων χρησιμοποιώντας R, Python ή άλλο. Παρέχετε μια εφαρμογή για όλους τους μαθητές εάν χρειάζεται (εάν οι μαθητές δεν μπορούν να αναπτύξουν τη δική τους σε σύντομο χρόνο).

4. Μαθηματική εξερεύνηση
   4.1. Εμπλέξτε τους μαθητές σε πρακτικές δραστηριότητες εξερευνώντας τελεστές συνόλων και διανυσματική άλγεβρα για να υπολογίσετε την ομοιότητα μεταξύ εγγράφων κειμένου σε ένα σώμα κειμένων.
   4.2. Διευκολύνετε τη συζήτησης σχετικά με τις μαθηματικές αρχές πίσω από την επεξεργασία κειμένου.

5. Ολοκληρωμένα projects
   5.1. Σχεδιάστε το project (καλύτερη λειτουργία ομοιότητας για έγγραφα κειμένου) και συντάξτε τον οδηγό.
   5.2. Συγκεντρώστε και σχολιάστε ένα σώμα κειμένων και ένα σύνολο ερωτημάτων για να αξιολογήσετε την ακρίβεια και την ανάκληση που παράγεται από τις υλοποιήσεις των συναρτήσεων ομοιότητας κειμένου από τους μαθητές. Η ακρίβεια και η ανάκληση θα αξιολογηθούν από τα στατικά ερωτήματα/έγγραφα δοκιμής (κάθε έγγραφο στο σώμα κειμένων θα σχολιαστεί με την κατάταξη ομοιότητας για κάθε μία από τις δοκιμαστικές περιπτώσεις), θα κρατήσει ένα σύνολο επικύρωσης (σχολιασμένο επίσης, αυτό μπορεί να είναι δύο ή τρία έγγραφα κειμένου, το καθένα με μερικούς όρους σαν να ήταν ένα ερώτημα για μια μηχανή αναζήτησης, για κάθε ένα από αυτά τα "ερωτήματα" δώστε την κατάταξη των εγγράφων στους μαθητές, αυτό θα χρησιμοποιηθεί για τον υπολογισμό της F1 στο τέλος και την ανακοίνωση του νικητή).
   5.3. Δώστε στους μαθητές πληροφορίες σχετικά με τις τεχνικές προ-επεξεργασίας (stop-word removal, stemming, etc.) και την υλοποίησή τους χρησιμοποιώντας τις βιβλιοθήκες εξόρυξης κειμένου που έχουν επιλεγεί για το έργο.

6. Σύνδεση με μηχανές αναζητήσης
   6.1. Προετοιμάστε το σενάριο επικύρωσης. Ένα "έγγραφο" μπορεί να είναι ένα ερώτημα όπως αυτά που χρησιμοποιούμε σε μηχανές αναζήτησης όπως το Google, όπως το "camélias porto" ή το "computer vision". Εμφάνιση σε πραγματικό χρησιμοποιώντας το Google. Δώστε στους μαθητές το σύνολο επικύρωσης, δηλαδή τρία έγγραφα, το καθένα με λίγες λέξεις-κλειδιά, σαν να ήταν ένα ερώτημα για μια μηχανή αναζήτησης. Εκτελέστε τις λειτουργίες ομοιότητας για να δώσετε μια

	κατάταξη των εγγράφων στο σώμα κειμένων και να δώσετε στους μαθητές την καλύτερη κατάταξη. Υπολογίστε τον νικητή χρησιμοποιώντας τη μέτρηση F1. Εξηγήστε στους μαθητές τι είναι το F1, η ανάκληση και η ακρίβεια. **Αξιολόγηση και προβληματισμός** 1. Αξιολογήστε την κατανόηση και την εφαρμογή των τελεστών και εννοιών της άλγεβρας συνόλων και διανυσμάτων από τους μαθητές μέσω αξιολογήσεων, παρουσιάσεων και γραπτών προβληματισμών βάσει εργασιών. 2. Ενθαρρύνετε τους μαθητές να προβληματιστούν σχετικά με τις μαθησιακές τους εμπειρίες, τονίζοντας τη σχέση μεταξύ μαθηματικών και εξόρυξης κειμένου. Ζητήστε από τους μαθητές να σχεδιάσουν και να παρουσιάσουν μια μεθοδολογία μηχανής αναζήτησης και ένα μη λειτουργικό πρωτότυπο χρησιμοποιώντας αυτά που έχουν μάθει.

[*] *υπό ανάπτυξη τα τελικά στοιχεία του πλαισίου*

3. Στόχοι και μεθοδολογίες

Μαθησιακοί σκοποί και στόχοι	1. Κατανόηση των γενικών εννοιών και τεχνικών μοντελοποίησης και επεξεργασίας που χρησιμοποιούνται στην εξόρυξη κειμένων 2. Εξερεύνηση των διεπιστημονικών συνδέσεων μεταξύ εξόρυξης κειμένου, μηχανών αναζήτησης και διανυσματικής άλγεβρας 3. Επεξήγηση των ομοιοτήτων μεταξύ εγγράφων κειμένου και άλλων μη δομημένων συνόλων δεδομένων ως εφαρμογές της διανυσματικής άλγεβρας
Μαθησιακά αποτελέσματα και αναμενόμενα αποτελέσματα	**Μαθησιακά Αποτελέσματα** A. Συζήτηση για θέματα υψηλού επιπέδου που σχετίζονται με τους τομείς της εξόρυξης κειμένων και των μηχανών αναζήτησης B. Περιγραφή της σχέσης μεταξύ διανυσματικής άλγεβρας, θεωρίας συνόλων και εξόρυξης κειμένου C. Εφαρμογή βασικών τεχνικών εξόρυξης κειμένου για την αντιμετώπιση απλών περιπτώσεων χρήσης **Αναμενόμενα Αποτελέσματα** 1. Λειτουργία ομοιότητας εγγράφου κειμένου σε R, Python ή άλλη γλώσσα προγραμματισμού
Προηγούμενες γνώσεις και Προαπαιτούμενα	1. Βασικές γνώσεις διανυσματικής άλγεβρας 2. Εξοικείωση με το Excel 3. Βασικές δεξιότητες προγραμματισμού λογισμικού 4. Άριστη χρήση εργαλείων πληροφορικής

Κίνητρα, Μεθοδολογία, Στρατηγικές, Υποστήριξη	1. Ανάθεση μαθητών σε μικρές ομάδες (3 ή 4 μαθητές). 2. Σχεδιάστε μια λύση, εφαρμόστε, δοκιμάστε και βελτιώστε με επαναληπτικό τρόπο. Χρησιμοποιήστε μια επαναληπτική μεθοδολογία ανάπτυξης. 3. Επισημάνετε τις συνδέσεις μεταξύ διανυσματικής άλγεβρας, μοντέλων εγγράφων, συνημίτονου και ομοιότητας. 4. Εξερευνήστε τις μηχανές αναζήτησης για να παρουσιάσετε τις σχέσεις μεταξύ της ομοιότητας κειμένου και των αποτελεσμάτων των μηχανών αναζήτησης. 5. Καθοδηγήστε τους μαθητές σε μια εξελικτική πορεία από τα πιο απλά μοντέλα (Boolean) σε πιο σύνθετες προσεγγίσεις (TFxIDF), εισάγοντας προκλήσεις βήμα προς βήμα ενώ δοκιμάζετε στο Excel με βασικές υλοποιήσεις σε μικρά έγγραφα με μία ή δύο σύντομες προτάσεις και λίγους διακριτούς όρους, από ένα λεξικό 10 ή 20 όρων.

4. Προετοιμασία και μέσα

Προετοιμασία, Διευθέτηση χώρου, *Συμβουλές αντιμετώπισης προβλημάτων*	Το εργαστήριο θα πραγματοποιηθεί σε αίθουσα διδασκαλίας για περίπου 20 μαθητές, σε ομάδες των 3 έως 4 μαθητών. Ιδανικά, η διάταξη της τάξης θα οργανωθεί σε 5 έως 7 ομάδες τραπεζιών όπου οι μαθητές από κάθε ομάδα μπορούν να καθίσουν ο ένας απέναντι από τον άλλο. Η αίθουσα χρειάζεται ένα δοκάρι και έναν τοίχο για παρουσιάσεις από όλους και έναν λευκό πίνακα με στυλό για να συζητήσετε ιδέες.
Πόροι, Εργαλεία, Υλικά, Επισυναπτόμενα, Εξοπλισμός	Ένα αποθετήριο στο GDrive, το Teams, το Github ή οποιονδήποτε άλλο πάροχο πρέπει να προετοιμαστεί εκ των προτέρων με όλο το περιβάλλον προγραμματισμού (R, Python, ...) και τα σώματα κειμένων που απαιτούνται για τα εργαστήρια. Πρέπει να δώσετε στους μαθητές ένα έγγραφο για να καθοδηγήσει τους μαθητές σε όλο το μάθημα / εργαστήριο, εξηγώντας λεπτομέρειες, αναμενόμενα αποτελέσματα, αξιολόγηση και μαθησιακά αποτελέσματα ανά μάθημα.
Υγιεινή και Ασφάλεια	

5. Εφαρμογή

Μαθησιακές δραστηριότητες, Διαδικασίες, Αναστοχασμοί	**Δομή εργαστηρίου** 1. Εισαγωγή [1 δραστηριότητα, 16 ώρες] 1.1. Επισκόπηση Εξόρυξης Δεδομένων και Μηχανικής Μάθησης: ιστορική προοπτική, εργασίες/προβλήματα, εφαρμογές, μοντελοποίηση μη δομημένων δεδομένων όπως εικόνες και κείμενο). [16 ώρες]

2. Μοντελοποίηση κειμένου [8 δραστηριότητες, 52 ώρες]
   2.1. Συγκεντρώστε τα σώματα κειμένων που θα χρησιμοποιηθούν για τις ασκήσεις και την εργασία του project (πρέπει να έχετε μικρά έγγραφα με πολύ μειωμένο λεξικό, συμπεριλάβετε έγγραφα όπου το TF μπορεί να κάνει τη διαφορά σε σύγκριση με το μοντέλο Boolean). [16 ώρες]
   2.2. Μοντέλα για την αναπαράσταση κειμένου, παραδείγματα: Μοντέλο Boolean, TF, TFxIDF, άλλο. Ανατρέξτε στο Bag of Words. [8 ώρες]
   2.3. Περιγράψτε συντελεστές συνόλου συνάφειας (ένωση, διασταύρωση κ.λπ.). [4 ώρες]
   2.4. Εφαρμόστε διάφορες εκδόσεις ενός εργαλείου στο Excel για να υπολογίσετε την ομοιότητα μεταξύ δύο εγγράφων κειμένου με βάση τελεστές συνόλου ή / και διανυσματική άλγεβρα (εσωτερικό γινόμενο, συνημίτονο, τομή και ένωση συνόλων, ...). [4 ώρες]
   2.5. Εισαγωγή διανυσματικής άλγεβρας: εσωτερικό προϊόν και συνημίτονο. [4 ώρες]
   2.6. Δείξτε τον αντίκτυπο/συνάφεια της συχνότητας των όρων σε σύγκριση με το Boolean. Χρησιμοποιήστε παραδείγματα. Καταστήστε σαφές ότι το συνημίτονο λειτουργεί καλά. [4 ώρες]
   2.7. Περιγράψτε τις προκλήσεις της ομοιότητας όταν βασίζεται μόνο στο TF (όροι που υπάρχουν σε όλα τα έγγραφα δεν έχουν διακριτική ισχύ), εναλλακτικές λύσεις. Χρησιμοποιήστε παραδείγματα. [4 ώρες]
   2.8. Παρουσιάστε το IDF και το TFxIDF, συζητήστε, σχεδιάστε και υλοποιήστε στο Excel ένα μέτρο ομοιότητας βασισμένο στο μοντέλο TFxIDF που θα χρησιμοποιηθεί για μικρά έγγραφα με μειωμένο λεξικό. [8 ώρες]

3. Εφαρμογή [2 δραστηριότητες, 32 ώρες]
   3.1. Ερευνήστε R, Python ή άλλες βιβλιοθήκες για εξόρυξη κειμένου. Προετοιμάστε ένα αποθετήριο και έναν οδηγό εγκατάστασης για τους μαθητές ώστε να μπορέσουν να εγκαταστήσουν το συγκεκριμένο περιβάλλον ανάπτυξης. [16 ώρες]
   3.2. Υλοποιήστε μια συνάρτηση για να υπολογίσετε την ομοιότητα μεταξύ δύο εγγράφων χρησιμοποιώντας R, Python ή άλλο. Παρέχετε μια εφαρμογή για όλους τους μαθητές εάν χρειάζεται (εάν οι μαθητές δεν μπορούν να αναπτύξουν τη δική τους σε σύντομο χρονικό διάστημα). [16 ώρες]

4. Μαθηματική εξερεύνηση [2 δραστηριότητες, 4 ώρες]
   4.1. Εμπλέξτε τους μαθητές σε πρακτικές δραστηριότητες εξερευνώντας τελεστές συνόλων και διανυσματική άλγεβρα για να υπολογίσετε την ομοιότητα μεταξύ εγγράφων κειμένου σε ένα σώμα. [2 ώρες]
   4.2. Διευκολύνετε τη συζήτηση σχετικά με τις μαθηματικές αρχές πίσω από την επεξεργασία κειμένου. [2 ώρες]

5. Ολοκληρωμένα projects [3 δραστηριότητες, 20 ώρες]
   5.1. Σχεδιάστε το project (καλύτερη λειτουργία ομοιότητας για έγγραφα κειμένου) και συντάξτε τον οδηγό. [8 ώρες]
   5.2. Συγκεντρώστε και σχολιάστε ένα σώμα κειμένων και ένα σύνολο ερωτημάτων για να αξιολογήσετε την ακρίβεια και την ανάκληση που

παράγεται από τις υλοποιήσεις των συναρτήσεων ομοιότητας κειμένου από τους μαθητές. Η ακρίβεια και η ανάκληση θα αξιολογηθούν από τα στατικά ερωτήματα/έγγραφα δοκιμής (κάθε έγγραφο στο σώμα κειμένων θα σχολιαστεί με την κατάταξη ομοιότητας για κάθε μία από τις δοκιμαστικές περιπτώσεις), θα κρατήσει ένα σύνολο επικύρωσης (σχολιασμένο επίσης, αυτό μπορεί να είναι δύο ή τρία έγγραφα κειμένου, το καθένα με μερικούς όρους σαν να ήταν ένα ερώτημα για μια μηχανή αναζήτησης. για κάθε ένα από αυτά τα "ερωτήματα" δώστε την κατάταξη των εγγράφων στους μαθητές, αυτό θα χρησιμοποιηθεί για τον υπολογισμό της F1 στο τέλος και την ανακοίνωση του νικητή). [8 ώρες]

5.3. Δώστε πληροφορίες σχετικά με τις τεχνικές προεπεξεργασίας (stop-word removal, stemming, etc.) και την υλοποίησή τους χρησιμοποιώντας τις βιβλιοθήκες εξόρυξης κειμένου που έχουν επιλεγεί για το project. [4 ώρες]

6. Σύνδεση με τις μηχανές αναζήτησης [1 δραστηριότητα, 8 ώρες]

6.1. Προετοιμάστε το σενάριο επικύρωσης. Ένα "έγγραφο" μπορεί να είναι ένα ερώτημα όπως αυτά που χρησιμοποιούμε σε μηχανές αναζήτησης όπως το Google, όπως το "camélias porto" ή το "computer vision". Εμφάνιση σε πραγματική χρήση με το Google. Δώστε στους μαθητές το σύνολο επικύρωσης, δηλαδή τρία έγγραφα, το καθένα με λίγες λέξεις-κλειδιά, σαν να ήταν ένα ερώτημα για μια μηχανή αναζήτησης. Εκτελέστε τις λειτουργίες ομοιότητας για να δώσετε μια κατάταξη των εγγράφων στο σώμα κειμένων και να δώσετε στους μαθητές την καλύτερη κατάταξη. Υπολογίστε τον νικητή χρησιμοποιώντας τη μέτρηση F1. Εξηγήστε στους μαθητές τι είναι το F1, η ανάκληση και η ακρίβεια. [8 ώρες]

Αξιολόγηση

Αξιολόγηση και προβληματισμός

1. Αξιολογήστε την κατανόηση και την εφαρμογή των τελεστών και εννοιών της άλγεβρας συνόλων και διανυσμάτων από τους μαθητές μέσω αξιολογήσεων, παρουσιάσεων και γραπτών προβληματισμών βάσει εργασιών.
2. Ενθαρρύνετε τους μαθητές να προβληματιστούν σχετικά με τις μαθησιακές τους εμπειρίες, τονίζοντας τη σχέση μεταξύ μαθηματικών και εξόρυξης κειμένου.
3. Ζητήστε από τους μαθητές να σχεδιάσουν και να παρουσιάσουν μια μεθοδολογία μηχανής αναζήτησης και ένα μη λειτουργικό πρωτότυπο χρησιμοποιώντας αυτά που έχουν μάθει.

Παρουσίαση - Αναφορά - Διαμοιρασμός

1. Μια συνάρτηση που υπολογίζει την ομοιότητα μεταξύ δύο εγγράφων κειμένου. Μια παρουσίαση στο PowerPoint που περιγράφει μια μεθοδολογία και ένα μη λειτουργικό πρωτότυπο μιας νέας μηχανής αναζήτησης που προτείνεται από τους μαθητές χρησιμοποιώντας αυτό που έχει αναπτυχθεί στο εργαστήριο (τη λειτουργία ομοιότητας εγγράφου).

Επεκτάσεις - Άλλες πληροφορίες

STEAME ACADEMY Πρωτότυπο/Οδηγός για τη μάθηση και τη δημιουργικότητα Προσέγγιση
Διαμόρφωση σχεδίου δράσης

Σημαντικά βήματα της μαθησιακής προσέγγισης STEAME:

ΣΤΑΔΙΟ Ι: Προετοιμασία από έναν ή περισσότερους εκπαιδευτικούς

1. Διατύπωση αρχικών σκέψεων σχετικά με τους θεματικούς τομείς/περιοχές που πρέπει να καλυφθούν
2. Σύνδεση με την πραγματικότητα του ευρύτερου περιβάλλοντος / εργασία / επιχειρήσεις / γονείς / κοινωνία / περιβάλλον / ηθική
3. Ηλικιακή ομάδα-στόχος των μαθητών - Σύνδεση με το επίσημο πρόγραμμα σπουδών - Καθορισμός στόχων και σκοπών
4. Οργάνωση των εργασιών των εμπλεκόμενων μερών - Ορισμός Συντονιστή - Χώροι εργασίας κ.λπ.

ΣΤΑΔΙΟ ΙΙ: Διαμόρφωση σχεδίου δράσης (Βήματα 1-18)

Προετοιμασία (από τους εκπαιδευτικούς)

1. Σύνδεση με τον πραγματικό κόσμο - Αναστοχασμός
2. Κινητοποίηση - Κίνητρα
3. Διατύπωση ενός προβλήματος (ενδεχομένως σε στάδια ή φάσεις) που προκύπτει από τα παραπάνω

Ανάπτυξη (από τους μαθητές) - Καθοδήγηση και αξιολόγηση (στο 9-11, από τους εκπαιδευτικούς)

4. Ανάπτυξη υποβάθρου- Αναζήτηση / Συγκέντρωση πληροφοριών
5. Απλοποίηση του ζητήματος - Διαμόρφωση του προβλήματος σε έναν περιορισμένο αριθμό απαιτήσεων
6. Διατύπωση περίπτωσης - Σχεδιασμός - προσδιορισμός υλικών για οικοδόμηση / ανάπτυξη / δημιουργία
7. Κατασκευή - Ροή εργασιών – εφαρμογή των project
8. Παρατήρηση -Πειραματισμός - Αρχικά συμπεράσματα
9. Τεκμηρίωση - Αναζήτηση θεματικών περιοχών (πεδία ΤΝ) που σχετίζονται με το υπό μελέτη θέμα - Επεξήγηση με βάση τις υπάρχουσες θεωρίες ή/και τα εμπειρικά αποτελέσματα
10. Συγκέντρωση αποτελεσμάτων/πληροφοριών με βάση τα σημεία 7, 8, 9
11. Πρώτη ομαδική παρουσίαση από τους μαθητές

Διαμόρφωση και αποτελέσματα (από τους μαθητές) - Καθοδήγηση και αξιολόγηση (από τους εκπαιδευτικούς)

12. Διαμόρφωση μοντέλων STEAME για την περιγραφή / αναπαράσταση / απεικόνιση των αποτελεσμάτων
13. Μελέτη των αποτελεσμάτων στο σημείο 9 και εξαγωγή συμπερασμάτων, χρησιμοποιώντας το σημείο 12

14. Εφαρμογές στην καθημερινή ζωή - Προτάσεις για την ανάπτυξη του σημείου 9 (Επιχειρηματικότητα - SIL ημέρες)

Ανασκόπηση (από εκπαιδευτικούς)

15. Επανεξέταση του προβλήματος και επανεξέταση και υπό περισσότερο απαιτητικές συνθήκες

Ολοκλήρωση έργου (από τους μαθητές) - Καθοδήγηση και αξιολόγηση (από τους εκπαιδευτικούς)

16. Επανάληψη των βημάτων 5 έως 11 με πρόσθετες ή νέες απαιτήσεις όπως διατυπώνονται στο σημείο 15
17. Διερεύνηση - Μελέτες περιπτώσεων - Επέκταση - Νέες θεωρίες - Έλεγχος νέων συμπερασμάτων
18. Παρουσίαση συμπερασμάτων - Τακτικές επικοινωνίας.

ΣΤΑΔΙΟ ΙΙΙ: STEAME ACADEMY Δράσεις και συνεργασία σε δημιουργικά έργα για μαθητές σχολείων

Τίτλος του έργου: _____
Σύντομη περιγραφή/προγραμματισμός των οργανωτικών διευθετήσεων/αρμοδιοτήτων για δράση

ΣΤΑΔΙΟ	**Δραστηριότητες/Βήματα** Εκπαιδευτικός 1(Τ1) Συνεργασία με τον Τ2 και καθοδήγηση των μαθητών	**Δραστηριότητες /Βήματα** **Από Μαθητές** Ηλικιακή ομάδα: ____	**Δραστηριότητες /Βήματα** Εκπαιδευτικός 2 (Τ2) Συνεργασία με Τ1 και καθοδήγηση των μαθητών
A	Προετοιμασία των σταδίων 1,2,3		Συνεργασία στο βήμα 3
B	Καθοδήγηση στο βήμα 9	4,5,6,7,8,9,10	Καθοδήγηση και υποστήριξη στο βήμα 9
C	Δημιουργική αξιολόγηση	11	Δημιουργική αξιολόγηση
D	Καθοδήγηση	12	Καθοδήγηση
E	Καθοδήγηση	13 (9+12)	Καθοδήγηση
F	Οργάνωση (SIL) STEAME στην καθημερινή ζωή	14 Συνάντηση με εκπροσώπους φορέων	Οργάνωση (SIL) STEAME στην καθημερινή ζωή
G	Προετοιμασία του βήματος 15		Συνεργασία στο βήμα 15
H	Καθοδήγηση	16 (επανάληψη 5-11)	Καθοδήγηση και υποστήριξη
I	Καθοδήγηση	17	Καθοδήγηση και υποστήριξη
K	Δημιουργική αξιολόγηση	18	Δημιουργική αξιολόγηση