

LESSON PLAN

ΑΣΦΑΛΕΙΑ ΣΤΗΝ ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ

ΣΤΟΧΕΥΜΕΝΟ ΚΟΙΝΟ

Μαθητές Ε και ΣΤ Δημοτικού.

ΔΙΑΡΚΕΙΑ ΥΛΟΠΟΙΗΣΗΣ

2 διδακτικές ώρες (μαζί με το φύλλο εργασίας)

ΔΙΔΑΚΤΙΚΟΙ ΣΤΟΧΟΙ ΜΑΘΗΜΑΤΟΣ

- ✓ Να κατανοήσουν τι είναι η Τεχνητή Νοημοσύνη (ΤΝ ή AI-Artificial Intelligence) και πώς λειτουργεί ως εργαλείο.
- ✓ Να αναγνωρίζουν τη σημασία της προστασίας των προσωπικών δεδομένων όταν χρησιμοποιούν εφαρμογές ΤΝ.
- ✓ Να εντοπίζουν πιθανά ψεύτικο ψηφιακό περιεχόμενο (deepfakes, παραπληροφόρηση) και να αναπτύξουν κριτική σκέψη.
- ✓ Να κατανοήσουν την έννοια της προκατάληψης (bias) στην εκπαίδευση της ΤΝ και τη σημασία της ποικιλομορφίας στα δεδομένα.
- ✓ Να συνειδητοποιήσουν ότι οι άνθρωποι έχουν την ευθύνη για τις αποφάσεις που λαμβάνονται με τη βοήθεια της ΤΝ.
- ✓ Να νιώσουν αυτοπεποίθηση ότι μπορούν να χρησιμοποιούν εφαρμογές τεχνητής νοημοσύνης με ασφάλεια και υπευθυνότητα.
- ✓ Να αναπτύξουν μια στάση κριτικής σκέψης απέναντι σε πληροφορίες που προέρχονται από ψηφιακές πηγές.
- ✓ Να συνεργάζονται σε μικρές ομάδες και να εκφράζουν επιχειρήματα σχετικά με την υπεύθυνη χρήση της ΤΝ.



ΑΝΑΛΥΤΙΚΗ ΠΟΡΕΙΑ ΔΙΔΑΣΚΑΛΙΑΣ

<p>Διαφάνειες 2 & 3</p> <p>Εισαγωγή στην ΤΝ</p> <p>Διάρκεια 5'</p>	<p>Ο στόχος των 2 πρώτων διαφανειών είναι να κατανοήσουν οι μαθητές και οι μαθήτριες ότι η τεχνητή νοημοσύνη είναι ένα εργαλείο που δεν σκέφτεται ή αισθάνεται όπως οι άνθρωποι. Ο/Η εκπαιδευτικός ξεκινάει τη συζήτηση διερευνώντας την πρότερη γνώση των παιδιών για την ΤΝ. Ρωτάει αν χρησιμοποιούν εφαρμογές ΤΝ και πώς. Τυχόν λανθασμένες αντιλήψεις δεν διορθώνονται σε αυτή τη φάση. Ο/Η εκπαιδευτικός τονίζει ότι η τεχνητή νοημοσύνη είναι ένα σύστημα με προηγμένες ειδικές ικανότητες. Έμφαση δίνεται στο ότι εμείς χρησιμοποιούμε την ΤΝ. Γι' αυτό είναι λάθος να πούμε «η ΤΝ μου είπε τι να γράψω στην έκθεση». Η αποφυγή ανθρωπομορφισμού στην ΤΝ είναι σημαντική γιατί όταν τα παιδιά αποδίδουν ανθρώπινες ιδιότητες στην ΤΝ, μπορεί να μπερδευτούν και να υπερτιμήσουν τις δυνατότητές της. Επίσης αν νομίζουν ότι η ΤΝ «αποφασίζει μόνη της», ίσως αποποιηθούν την ευθύνη για όσα κάνουν με τη βοήθειά της. Χρειάζεται να κατανοήσουν ότι οι τελικές αποφάσεις ανήκουν σε εμάς.</p>
<p>Διαφάνεια 4,5 & 6</p> <p>Πώς μαθαίνει η ΤΝ</p> <p>Διάρκεια: 5' λεπτά</p>	<p>Ο στόχος των διαφανειών 4, 5 και 6 είναι να κατανοήσουν οι μαθητές και οι μαθήτριες ότι τα συστήματα τεχνητής νοημοσύνης μπορεί να λειτουργούν με δύο τρόπους: είτε ακολουθώντας σταθερούς κανόνες, είτε μαθαίνοντας από δεδομένα που τους παρέχουμε. Τα συστήματα τεχνητής νοημοσύνης βασίζονται στην ανάλυση δεδομένων για να εντοπίζουν πρότυπα και να παράγουν προβλέψεις. Αν τα δεδομένα που χρησιμοποιούνται είναι ελλιπή ή περιέχουν σφάλματα, τότε τα αποτελέσματα της επεξεργασίας μπορεί να είναι ανακριβή ή παραπλανητικά.</p> <p>Ο/Η εκπαιδευτικός ξεκινάει τη συζήτηση εξηγώντας ότι κάποια παλαιότερα συστήματα λειτουργούσαν απλά με σταθερούς κανόνες, χωρίς να «μαθαίνουν» (π.χ. μια αριθμομηχανή). Σήμερα όμως οι εφαρμογές τεχνητής νοημοσύνης μαθαίνουν μέσα από τα δεδομένα εκπαίδευσης που συλλέγουν και αναπροσαρμόζουν τις προτάσεις ή τις ενέργειές τους ανάλογα με αυτά.</p> <p>Παράδειγμα: Αν ένα παιδί παρακολουθεί συνεχώς βίντεο με γάτες, το σύστημα τεχνητής νοημοσύνης θα εμφανίζει περισσότερα παρόμοια βίντεο, βασιζόμενο σε αυτό το πρότυπο χρήσης και χωρίς να λαμβάνει υπόψη ότι μπορεί να υπάρχουν και άλλες προτιμήσεις.</p> <p>Ο/Η εκπαιδευτικός εξηγεί ότι η τεχνητή νοημοσύνη δεν καταλαβαίνει ούτε κρίνει όπως οι άνθρωποι. Απλώς επεξεργάζεται τα δεδομένα που της δίνουμε και λειτουργεί σύμφωνα με συγκεκριμένους κανόνες (αλγόριθμους) που έχουν σχεδιάσει οι άνθρωποι. Αν τα δεδομένα είναι περιορισμένα ή λανθασμένα, το παραγόμενο αποτέλεσμα θα είναι επίσης ανακριβές ή παραπλανητικό.</p>



<p>Διαφάνεια 7,8 & 9</p> <p>Προστατεύω τα προσωπικά μου δεδομένα</p> <p>Διάρκεια: 15'</p>	<p>Ακολουθώντας, στις διαφάνειες 7 έως και 9, περνάμε στη σημασία της προστασίας των προσωπικών δεδομένων όταν χρησιμοποιούμε εφαρμογές TN.</p> <p>Ο/Η εκπαιδευτικός ρωτάει τα παιδιά αν έχουν χρησιμοποιήσει εφαρμογές που ζητούν προσωπικές πληροφορίες και τι είδους πληροφορίες ήταν αυτές. Τους εξηγεί ότι κάθε φορά που μοιραζόμαστε προσωπικά στοιχεία (όνομα, ηλικία, φωτογραφίες, προτιμήσεις), αυτά μπορεί να χρησιμοποιούνται για να "εκπαιδεύσουν" εφαρμογές TN και να τα μοιραστούν με άλλους χωρίς την άδειά μας.</p> <p>Μέσα από απλά σενάρια που παρουσιάζονται (π.χ. εγκατάσταση εφαρμογής αναγνώρισης ράτσας σκύλων ή προτάσεις μουσικής), τα παιδιά καλούνται να σκεφτούν πότε είναι ασφαλές και πότε όχι να δώσουν πληροφορίες.</p> <p>Δίνεται έμφαση στη λογική "Δίνω μόνο όσα είναι απολύτως απαραίτητα" και στη συμβουλή "Ρωτάω πάντα έναν ενήλικα πριν αποκαλύψω προσωπικά στοιχεία σε μια εφαρμογή."</p> <p>Τέλος, υπογραμμίζεται ότι η ευθύνη της προστασίας των δεδομένων μας ανήκει σε εμάς, και ότι πρέπει να είμαστε προσεκτικοί, ακριβώς γιατί η TN δεν καταλαβαίνει από μόνη της τι είναι ευαίσθητο ή τι πρέπει να προστατευτεί - τα αποτελέσματα που παράγει βασίζονται στα δεδομένα με τα οποία έχει εκπαιδευτεί.</p>
<p>Διαφάνειες 10-17</p> <p>Έλεγχος πληροφορίας</p> <p>Διάρκεια 15'</p>	<p>Ο στόχος των διαφανειών 10-17 είναι να κατανοήσουν οι μαθητές και οι μαθήτριες ότι η τεχνητή νοημοσύνη μπορεί να δημιουργεί περιεχόμενο που φαίνεται αληθινό αλλά δεν είναι, όπως ψεύτικες εικόνες, βίντεο (deepfakes) και ειδήσεις. Επιδιώκεται να αναπτύξουν δεξιότητες κριτικής αξιολόγησης ψηφιακού περιεχομένου και να μάθουν να αναγνωρίζουν ενδείξεις παραποίησης.</p> <p>Ο/Η εκπαιδευτικός ξεκινά τη συζήτηση εξηγώντας ότι η τεχνητή νοημοσύνη μπορεί να παραγάγει εικόνες, βίντεο ή κείμενα που προσομοιάζουν την πραγματικότητα, αλλά στην ουσία έχουν δημιουργηθεί ή τροποποιηθεί μέσω αλγοριθμικών διαδικασιών. Επισημαίνεται ότι το ψηφιακό περιεχόμενο στο διαδίκτυο δεν είναι πάντα αξιόπιστο και ότι απαιτείται προσεκτική παρατήρηση και κριτική αξιολόγηση για τον εντοπισμό ενδείξεων παραποίησης.</p> <p>Ο/Η εκπαιδευτικός παρουσιάζει με απλά λόγια τις βασικές ενδείξεις που μπορεί να δείχνουν ότι μια εικόνα ή ένα video είναι παραποιημένα. Δίνονται ως παράδειγμα 3 φωτογραφίες οι οποίες και οι τρεις είναι αποτέλεσμα τεχνητής νοημοσύνης και καλούνται τα παιδιά να αναγνωρίσουν αν είναι αληθινές ή παραποιημένες. Στην πρώτη φωτογραφία η παραποίηση είναι εμφανής, στη δεύτερη αποδεικνύεται από τα παραμορφωμένα χέρια και στην τρίτη από τα έντονα χρώματα, τις σκιάσεις και τα μαλλιά. Στη συνέχεια ο/η</p>

	<p>εκπαιδευτικός περνάει στην αναγνώριση ψεύτικων ειδήσεων και κειμένων. Δίνονται δυο παραδείγματα όπου οι μαθητές και οι μαθήτριες θα πρέπει να αποφασίσουν με μια πρώτη ματιά αν θα τα κοινοποιήσουν ή θα τα αγνοήσουν. Και στις δυο περιπτώσεις θα πρέπει να τα αγνοήσουν καθώς στο πρώτο παράδειγμα εντοπίζονται γραμματικά λάθη και η εικόνα είναι αποτέλεσμα επεξεργασίας ΤΝ. Στο δεύτερο παράδειγμα η πληροφορία δεν βασίζεται σε επιστημονικά δεδομένα, είναι υπερβολική και η εικόνα είναι αποτέλεσμα επεξεργασίας ΤΝ. Οι μαθητές/τριες θα πρέπει να συνειδητοποιήσουν ότι πάντα αν θέλουμε να κοινοποιήσουμε κάτι θα πρέπει να ελέγξουμε την αξιοπιστία του και από άλλες πηγές. Πάντα βασιζόμαστε στην κριτική μας σκέψη πριν κοινοποιήσουμε.</p>
<p>Διαφάνειες 18-22</p> <p>Προκατάληψη στην ΤΝ</p> <p>Διάρκεια 15'</p>	<p>Ο στόχος των διαφανειών 18–22 είναι να κατανοήσουν οι μαθητές και οι μαθήτριες ότι τα συστήματα τεχνητής νοημοσύνης βασίζονται σε δεδομένα για να λειτουργήσουν και ότι, όταν τα δεδομένα αυτά περιλαμβάνουν προκαταλήψεις ή δεν είναι επαρκώς ποικιλόμορφα και αντικειμενικά, τα παραγόμενα αποτελέσματα ενδέχεται να αντικατοπτρίζουν αυτές τις στρεβλώσεις. Στόχος είναι να αντιληφθούν πώς οι επιλογές στα δεδομένα επηρεάζουν την αξιοπιστία και τη δικαιοσύνη των εφαρμογών ΤΝ.</p> <p>Επιδιώκεται να αναπτύξουν την κριτική τους σκέψη ώστε να αναγνωρίζουν ότι τα αποτελέσματα που παράγει η τεχνητή νοημοσύνη δεν είναι πάντα αντικειμενικά και ότι η δίκαιη και σωστή χρήση της απαιτεί προσοχή στον τρόπο που σχεδιάζεται και στα δεδομένα που χρησιμοποιούνται κατά την εκπαίδευσή της.</p> <p>Ο/Η εκπαιδευτικός ξεκινάει τη συζήτηση εξηγώντας ότι η εκπαίδευση των συστημάτων ΤΝ βασίζεται σε μεγάλο πλήθος παραδειγμάτων που της παρέχονται. Αν αυτά τα παραδείγματα είναι περιορισμένα, τότε και τα παραγόμενα αποτελέσματα ενδέχεται να εμπεριέχουν προκαταλήψεις.</p> <p>Ο/Η εκπαιδευτικός παρουσιάζει απλά παραδείγματα:</p> <p>Αν χρησιμοποιηθούν μόνο φωτογραφίες αντρών γιατρών και γυναικών δασκάλων, όταν ζητηθεί από την ΤΝ να δημιουργήσει την εικόνα ενός γιατρού, πιθανότατα θα παραγάγει την εικόνα ενός άνδρα.</p> <p>Αν χρησιμοποιηθούν μόνο εικόνες άσπρων σκύλων, τότε η ΤΝ θα συσχετίζει την έννοια του σκύλου μόνο με το άσπρο χρώμα.</p> <p>Ακολουθούν δύο παραδείγματα όπου το αποτέλεσμα της τεχνητής νοημοσύνης - με βάση τα δεδομένα που έχουν χρησιμοποιηθεί για την εκπαίδευσή της - ενδέχεται να παρουσιάζει προκατάληψη (π.χ. μόνο τα αγόρια παίζουν ποδόσφαιρο, μόνο τα κορίτσια κάνουν μπαλέτο).</p> <p>Ο/Η εκπαιδευτικός τονίζει ότι για να είναι τα αποτελέσματα της Τεχνητής Νοημοσύνης αμερόληπτα, θα πρέπει να στηρίζονται σε μεγάλο πλήθος δεδομένων με ποικιλομορφία και αντικειμενικότητα. Χρησιμοποιεί τον παραλληλισμό ότι, όπως η</p>

	<p>μάθηση είναι πιο πλήρης όταν υπάρχει επαφή με ολόκληρη την πολυμορφία του κόσμου και όχι με ένα μόνο μέρος του, έτσι και η εκπαίδευση των συστημάτων ΤΝ απαιτεί ευρεία και αντιπροσωπευτική ποικιλία παραδειγμάτων, ανθρώπων, δραστηριοτήτων και εικόνων.</p> <p>Σε αυτό το σημείο οι μαθητές μπορούν να πειραματιστούν περισσότερο μέσω του εργαλείου Machine Learning for Kids ή το Teachable Machine Learning. Μπορούν να εκπαιδεύσουν το μοντέλο τους σωστά ώστε να μην παρουσιάζει προκατάληψη π.χ. να μην εισάγουν εικόνες μόνο λευκών ανθρώπων ώστε αν ζητήσουν από το μοντέλο να αναγνωρίσει έναν άνθρωπο διαφορετικού χρώματος, αυτό να μπορέσει να το κάνει.</p>
<p>Διαφάνειες 23-26</p> <p>Υπευθυνότητα στην ΤΝ</p> <p>Διάρκεια 10΄</p>	<p>Ο στόχος των διαφανειών 23–26 είναι να κατανοήσουν οι μαθητές και οι μαθήτριες ότι η τεχνητή νοημοσύνη μπορεί να χρησιμοποιείται ως εργαλείο υποστήριξης, αλλά την ευθύνη για τις αποφάσεις τη φέρουν πάντα οι άνθρωποι. Τα συστήματα ΤΝ δεν διαθέτουν επίγνωση της πραγματικότητας, συναισθήματα ή ικανότητα ηθικής κρίσης, επομένως απαιτείται συνειδητή και υπεύθυνη χρήση τους.</p> <p>Το ζητούμενο είναι να καλλιεργηθεί η κριτική στάση των παιδιών απέναντι σε αποφάσεις που βασίζονται σε εφαρμογές ΤΝ και να ενισχυθεί η κατανόηση ότι οι τελικές ευθύνες παραμένουν στους ανθρώπους.</p> <p>Ο/Η εκπαιδευτικός ξεκινάει τη συζήτηση εξηγώντας ότι τα συστήματα ΤΝ μπορούν να παρέχουν προτάσεις ή προβλέψεις, αλλά η λήψη αποφάσεων αποτελεί ανθρώπινη αρμοδιότητα.</p> <p>Είναι ένα εργαλείο υποστήριξης που δεν μπορεί να γνωρίζει την πλήρη εικόνα ή τα συναισθήματα των ανθρώπων.</p> <p>Ο/Η εκπαιδευτικός παρουσιάζει το παράδειγμα:</p> <p>«Φανταστείτε ότι έχουμε ένα ρομπότ στην αυλή του σχολείου και του λέμε: <i>Διάλεξε ποιος θα παίξει πρώτος στην τσουλήθρα</i>. Το ρομπότ βλέπει μόνο ποιος είναι πιο ψηλός ή ποιος έτρεξε πιο γρήγορα. Δεν καταλαβαίνει αν κάποιο παιδί έχει χτυπήσει, αν περιμένει πολλή ώρα, ή αν χθες δεν πρόλαβε να παίξει.»</p> <p>«Είναι αυτή μια δίκαιη απόφαση;» Το ρομπότ δεν μπορεί να καταλάβει τις ιδιαίτερες συνθήκες και συναισθήματα των παιδιών.</p> <p>Ο/Η εκπαιδευτικός εξηγεί ότι:</p> <p>Οι άνθρωποι μπορούν να σκεφτούν πιο σφαιρικά και δίκαια. Μπορούν να πάρουν αποφάσεις με βάση την ευγένεια, την αλληλεγγύη και τη δικαιοσύνη (π.χ. να δώσουν προτεραιότητα σε κάποιο παιδί που χθες δεν έπαιξε).</p> <p>Ο/Η εκπαιδευτικός συνοψίζει:</p> <p>«Η ΤΝ δεν έχει συναισθήματα, δεν καταλαβαίνει ανάγκες ή αδικίες. Μπορεί να μας βοηθήσει να βρούμε επιλογές, αλλά οι τελικές</p>

αποφάσεις πρέπει να τις παίρνουμε εμείς. Όπως ένα κομπιουτεράκι μας βοηθά στους υπολογισμούς, αλλά εμείς αποφασίζουμε τι να υπολογίσουμε, έτσι και η ΤΝ είναι βοηθός, όχι αρχηγός.»

