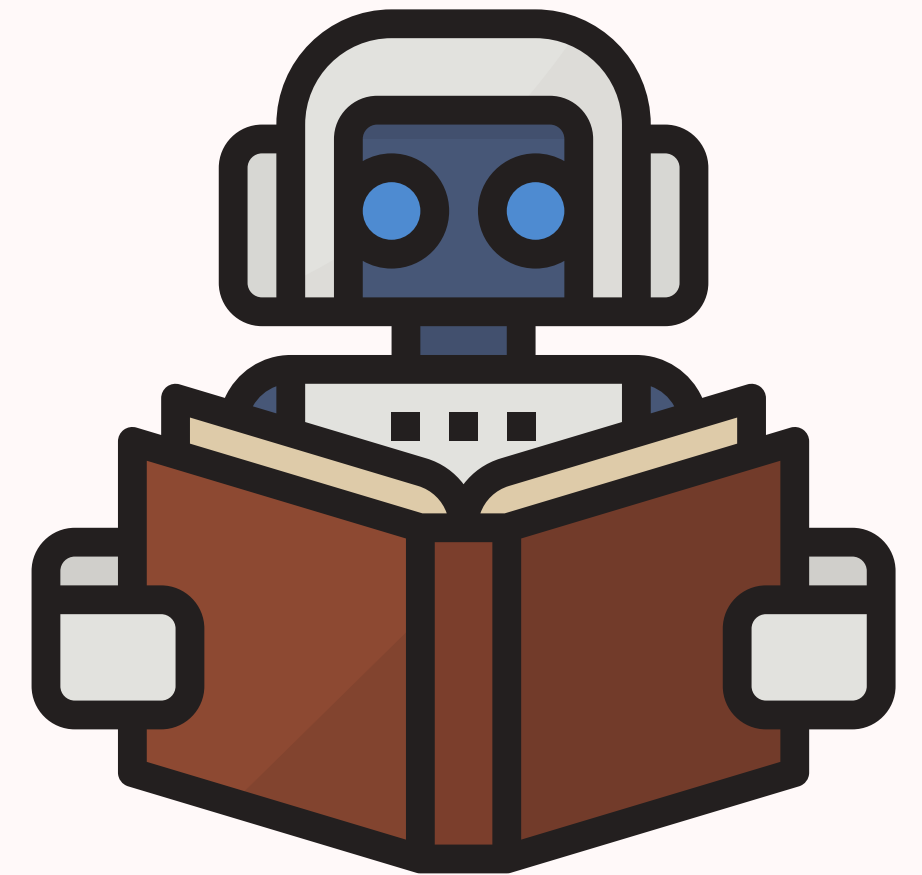


BIG DATA & MACHINE LEARNING

Βασίλειος Τίλλης

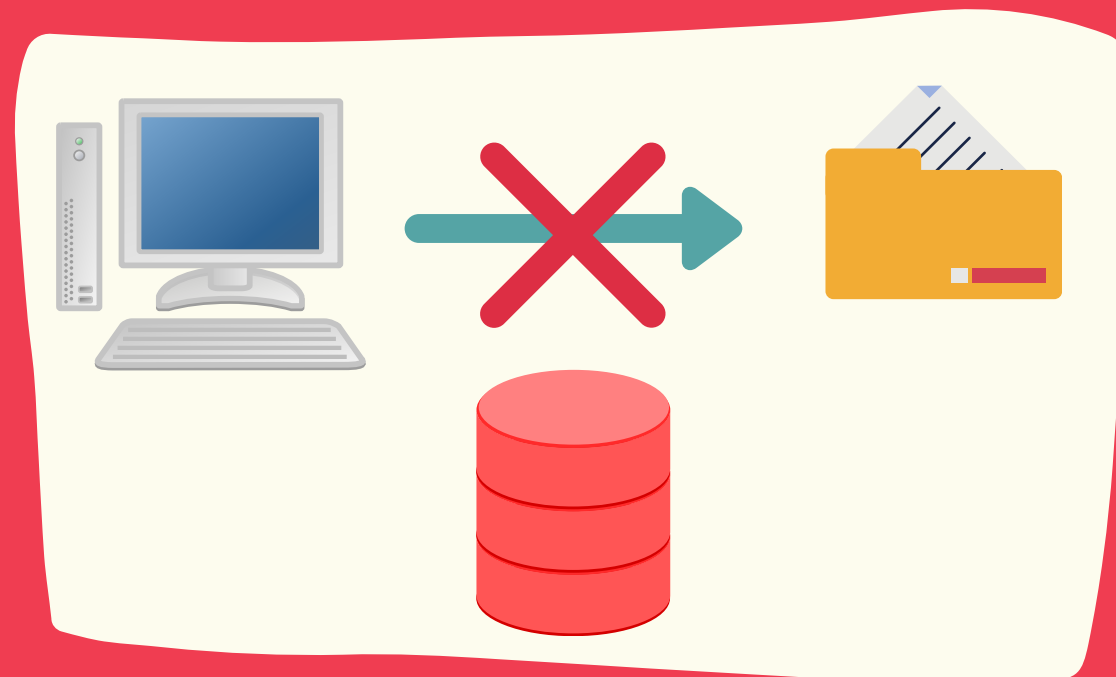


Τι είναι τα Big Data



Περίπου 40 Exabytes
δημιουργούνται κάθε μήνα
από έναν μόνο χρήστη Smartphone

40 Exabytes x 5 εκατομμύρια χρήστες Smartphone



**BIG
DATA**

Τα δεδομένα που παράγονται κάθε λεπτό στο Διαδίκτυο



3,8
εκατομμύρια
αναζητήσεις



1 εκατομμύριο
χρήστες
πραγματοποιούν
είσοδο (log in)



4,5
εκατομμύρια
παρακολουθήσεις
βίντεο



188 εκατομμύρια
e-mail



2,1
εκατομμύρια
snaps



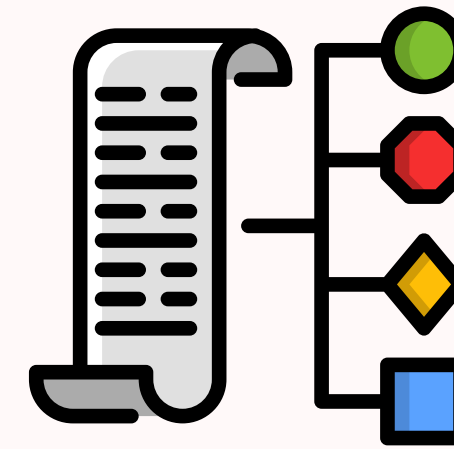
THE 5V'S TON BIG DATA



Volume



Velocity



Variety



Veracity

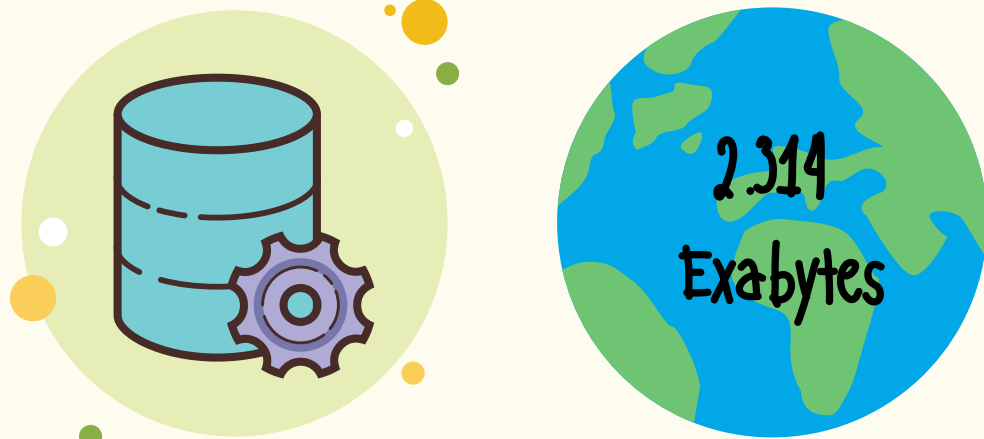


Value

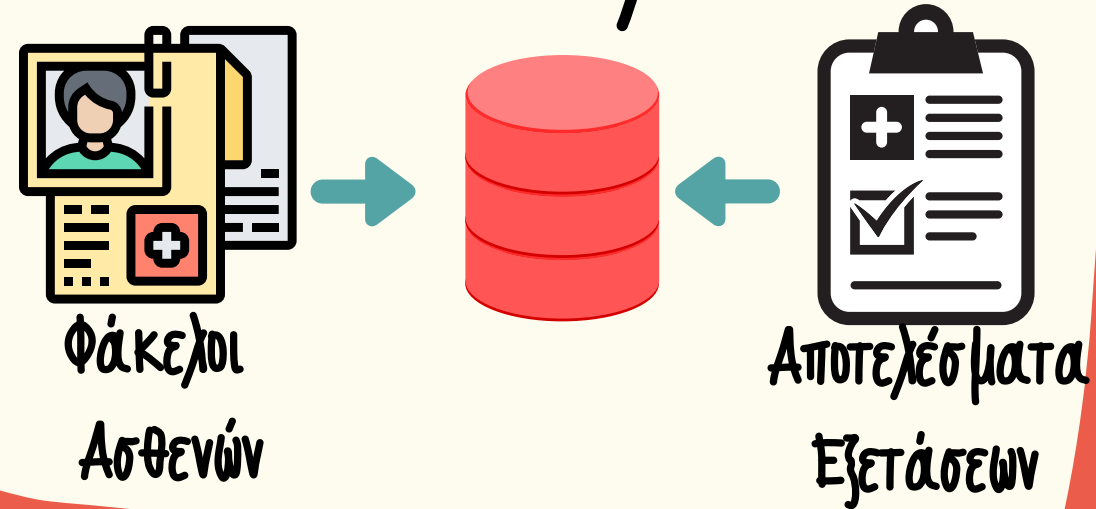


Παράδειγμα εφαρμογής Big Data στην Βιομηχανία της Υγείας

Volume



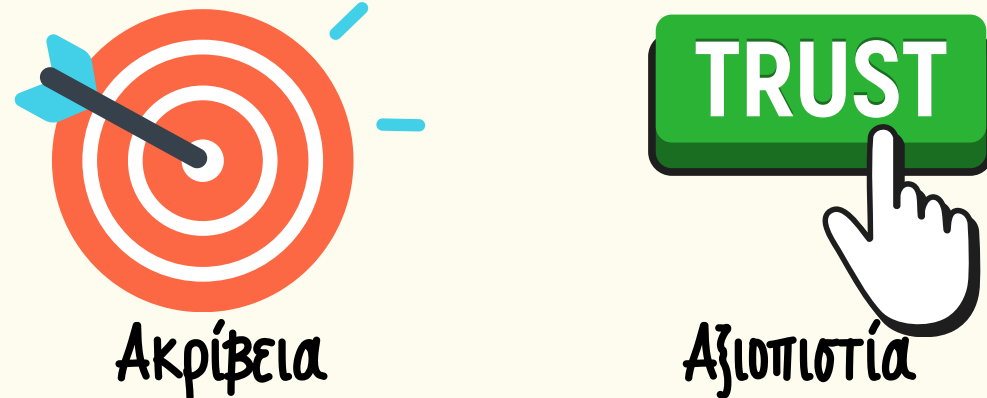
Velocity



Variety

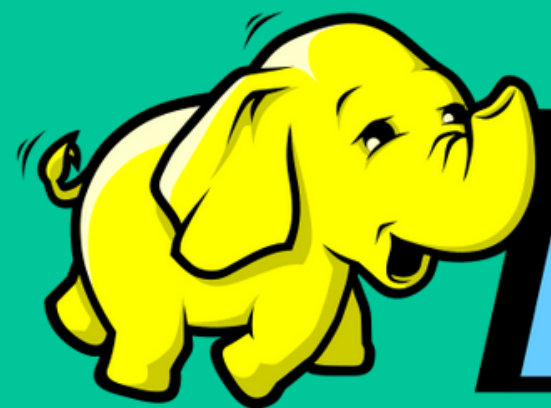


Veracity

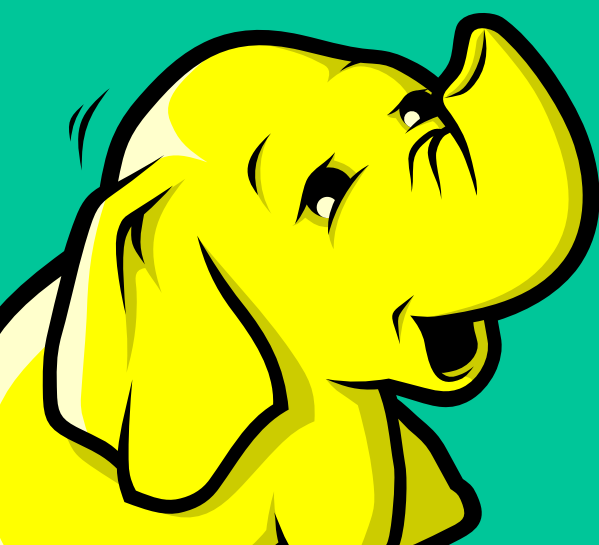
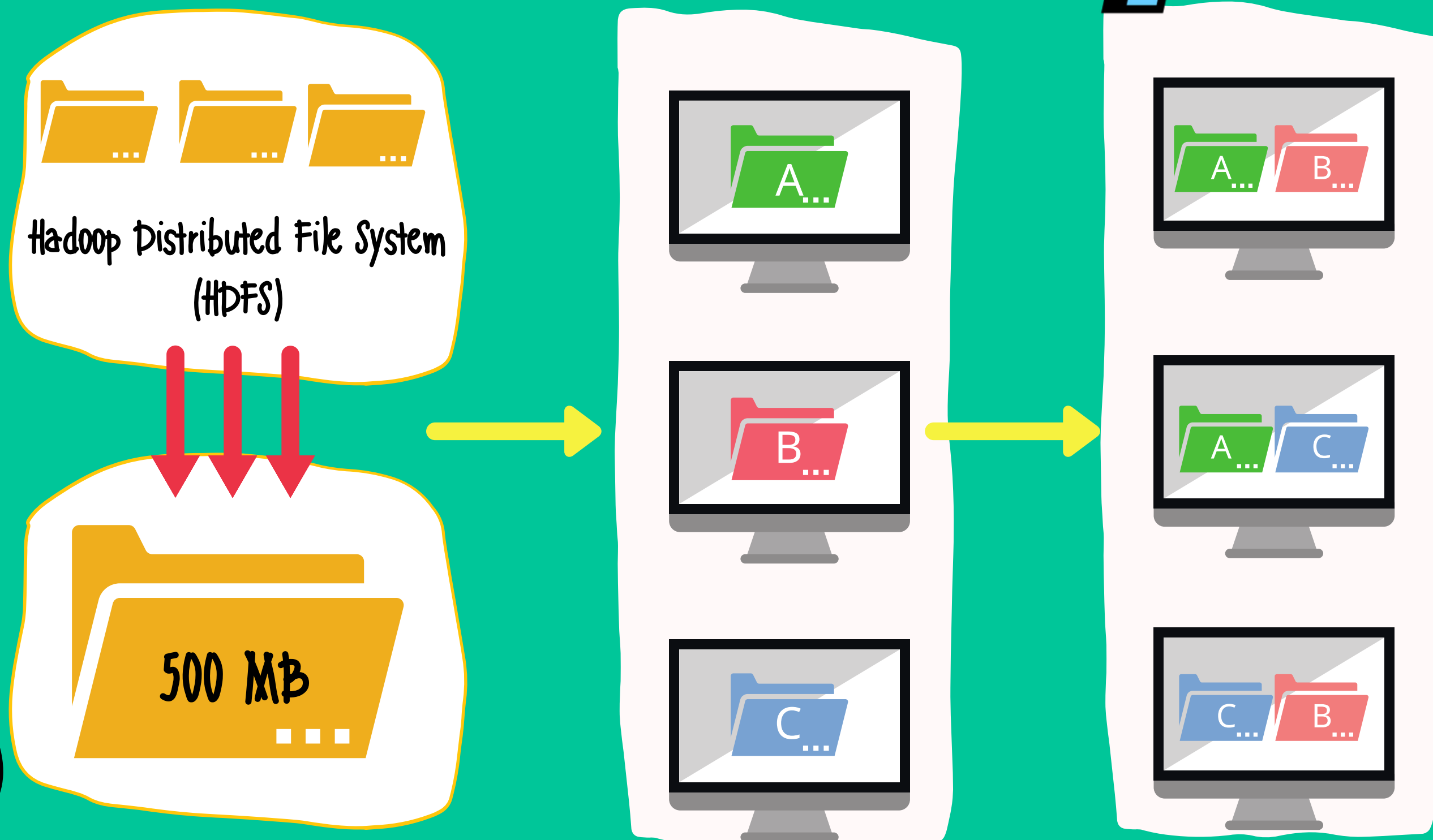


Value





hadoop





Διεργασία Β

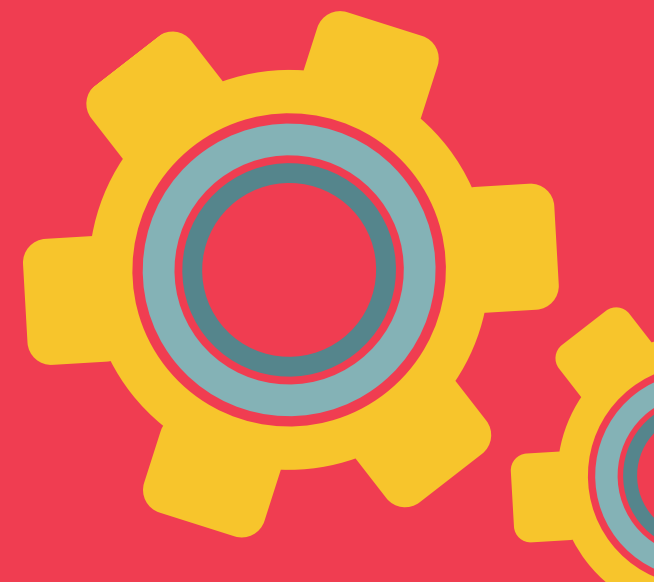


Διεργασία Τ



Διεργασία Δ

Parallel Processing



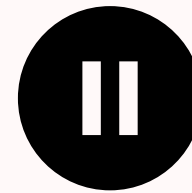
Εφαρμογή των Big Data στην καθημερινότητα



Χρήστες
Video Games



Σχεδιαστές
Video Games



Pause



Restart



Quit

Επίγνωση - Διορατικότητα
Insight



Βελτίωση
User Experience



Big Data



Τυφώνας
SANDY
2012



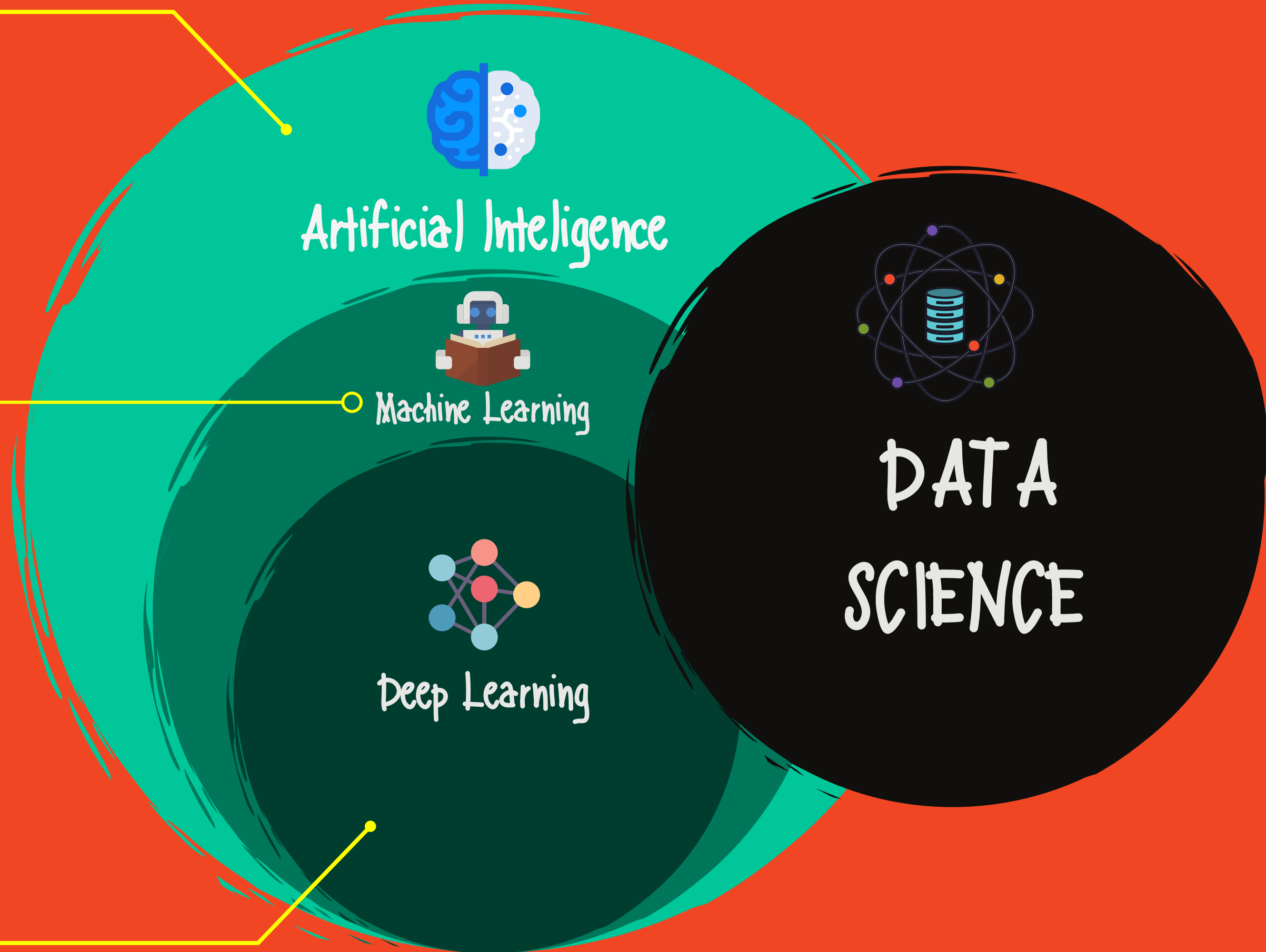
Πρόβλεψη πορείας
τυφώνων

MACHINE LEARNING - ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ

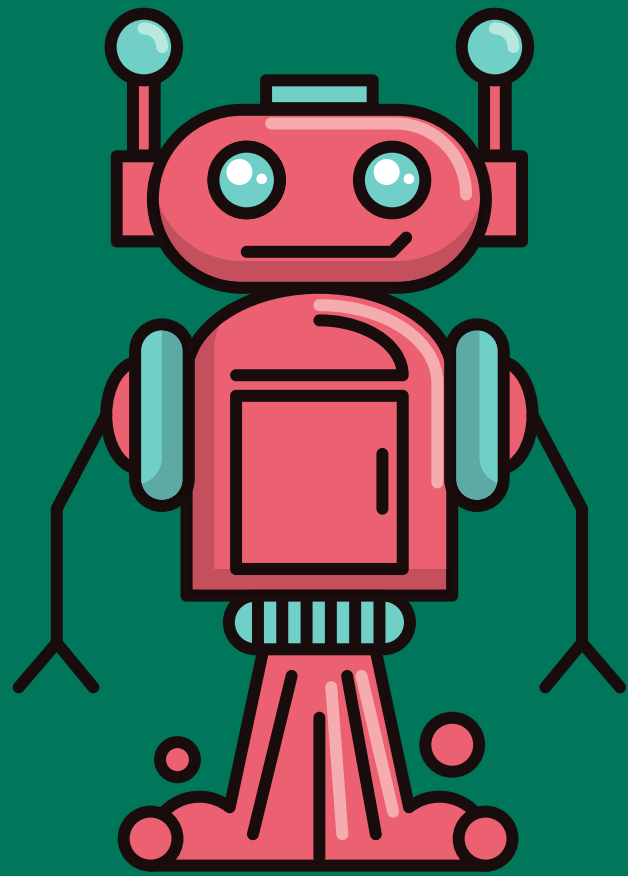
Η τεχνική που δίνει την δυνατότητα στις μηχανές να μιμηθούν μία ανθρώπινη συμπεριφορά

Θεωρείται ως μία υποκατηγορία της τεχνητής νοημοσύνης (A.I.) που εστιάζει στην κατασκευή εφαρμογών που μαθαίνουν από τα εισαγόμενα δεδομένα και βελτιώνουν την ακρίβειά τους με την πάροδο του χρόνου χωρίς να προγραμματίζονται να το κάνουν αυτό .

Η βαθιά μάθηση θεωρείται ως μία υποκατηγορία της Μηχανικής Μάθησης και συνδυάζει την πρόοδο της υπολογιστικής ισχύος και των ειδικών τύπων νευρωνικών δικτύων, ώστε να μαθαίνει περίπλοκα πρότυπα σε μεγάλες ποσότητες δεδομένων.

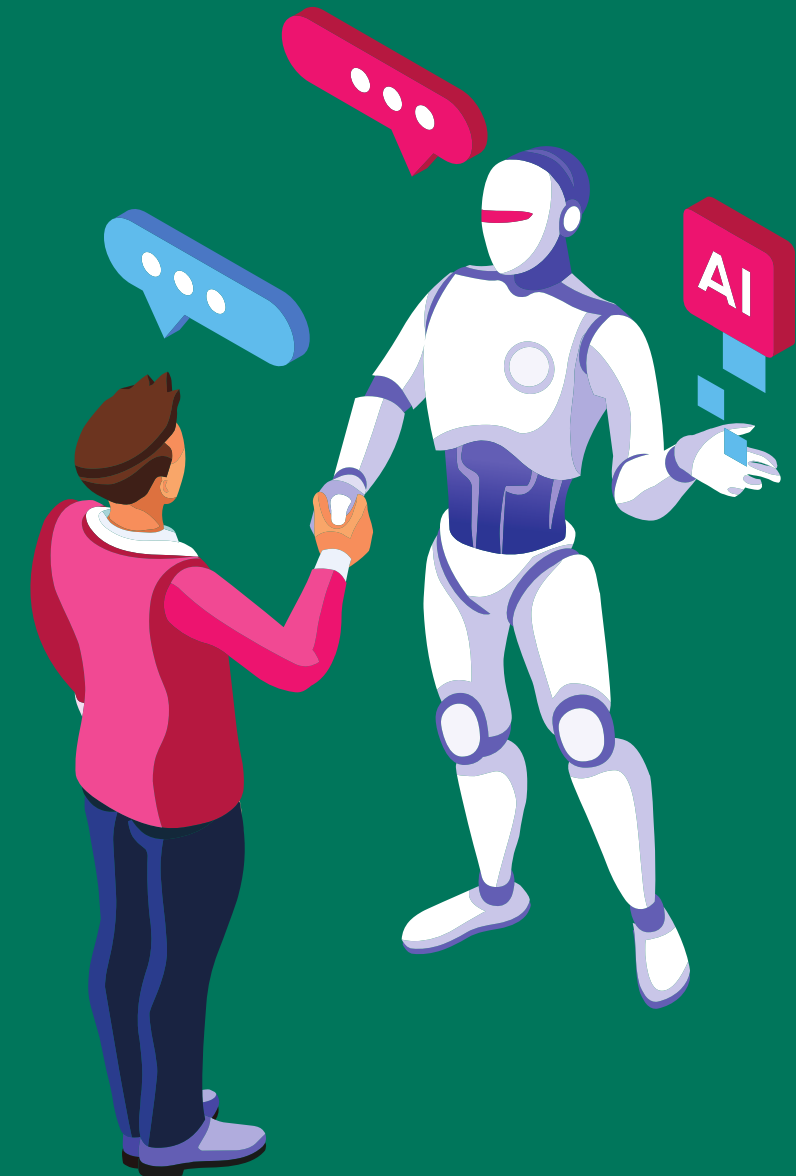


Τι είναι η Τεχνητή Νοημοσύνη - Artificial Intelligence (A.I.)



Με τον όρο A.I. εννοούμε μία μορφή ευφυίας που μπορεί να εμφανίσει μια υπολογιστική μηχανή.

Που σημαίνει την εκμάθηση, την κατανόηση και τη χρήση των γνώσεων που έχει μάθει ένα σύστημα Τεχνητής Νοημοσύνης για την εκτέλεση ενός ή περισσότερων εργασιών ή στόχων.



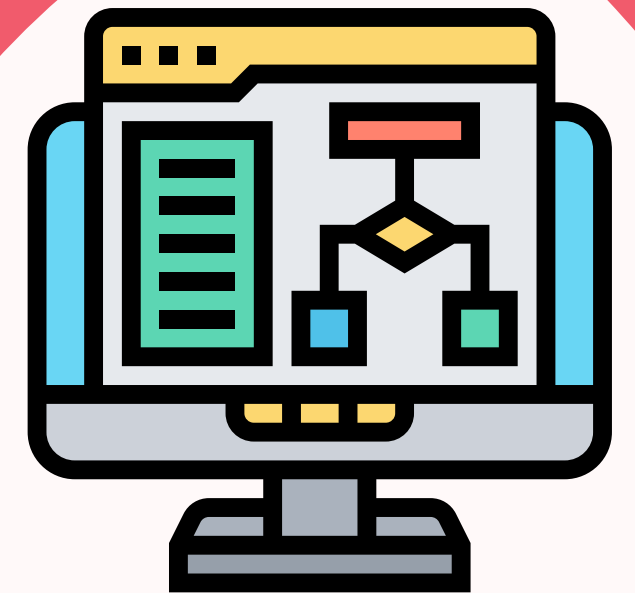
Πως λειτουργεί η Μηχανική Μάθηση - Machine Learning



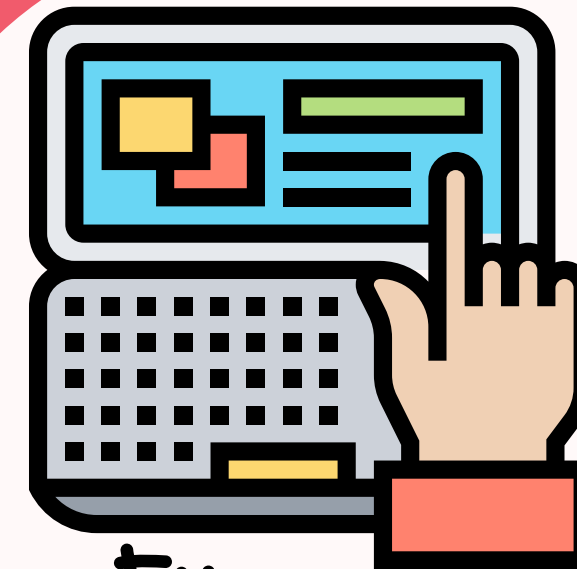
Είσοδος
Δεδομένων



Ανάλυση
Δεδομένων



Μοντέλο
Αλγόριθμου



Εύρεση
Μοτίβων



Προβλέψεις



Εφαρμογές Μηχανικής Μάθησης στην καθημερινότητα του ανθρώπου



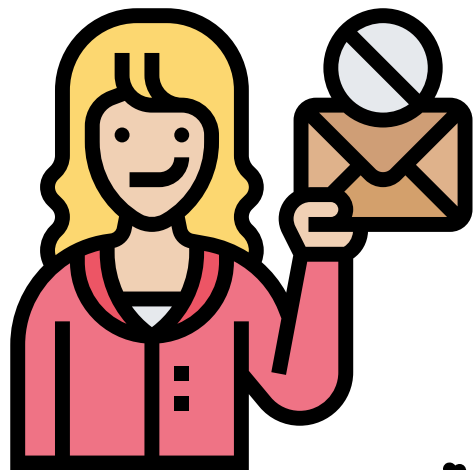
Image Recognition

Ένα μοντέλο μηχανικής μάθησης να παίρνει ως είσοδο δεδομένα με διάφορες εικόνες και να εκπαιδευτεί από εμάς, βάση τα δεδομένα που του έχουμε δώσει ως είσοδο, να αναγνωρίζει μόνο του νέες εικόνες.

Ένα μοντέλο μηχανικής μάθησης είναι ικανό να μεταφράσει την ομιλία σε κείμενο.



Speech Recognition

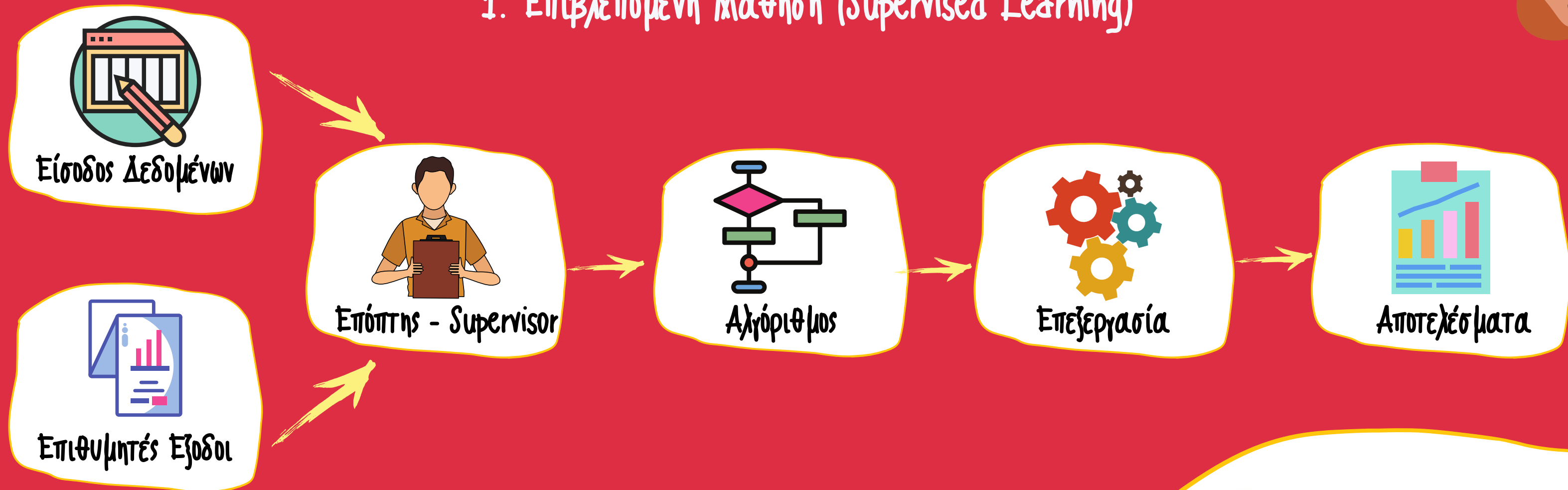


Ανεπιθύμητα Mail

Ο χρήστης επισημαίνοντας κάποια email ως ανεπιθύμητα, το μοντέλο θα μαθαίνει να απομνημονεύει όλα τα προηγούμενα email που φέρουν αυτή την ετικέτα (spam emails).

Μέθοδοι Μηχανικής Μάθησης

1. Επιβλεπόμενη Μάθηση (Supervised Learning)



Οι σημαντικότεροι αλγόριθμοι επιβλεπόμενης μάθησης περιλαμβάνουν εκείνους που χρησιμοποιούνται για Κατηγοριοποίηση (classification) και για Παλινδρόμηση (regression)



Regression

What is the temperature going to be tomorrow?

PREDICTION

84°

Fahrenheit °F -50 -40 -30 -20 -10 0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230



Classification

Will it be Cold or Hot tomorrow?

PREDICTION

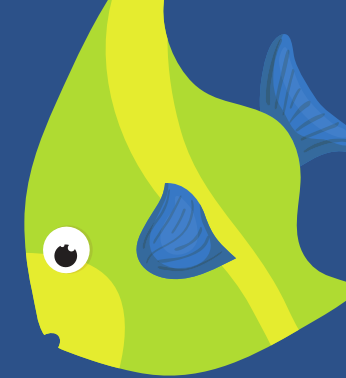
COLD

HOT

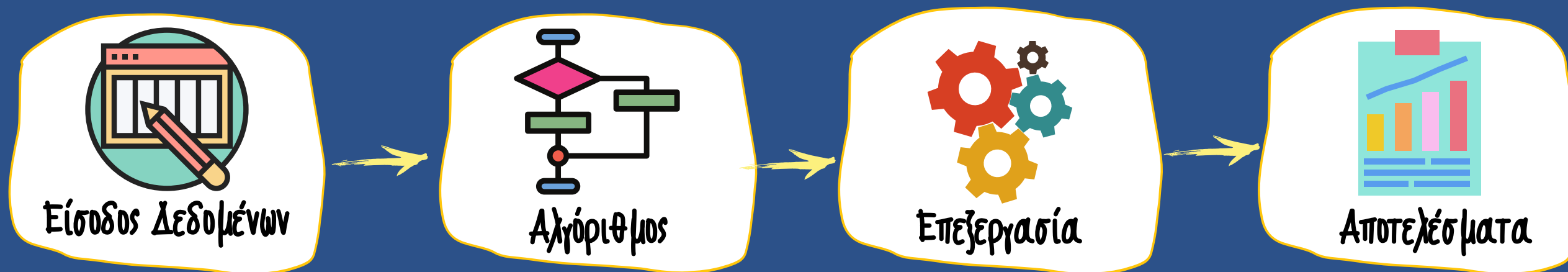
Fahrenheit °F -50 -40 -30 -20 -10 0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230



Μέθοδοι Μηχανικής Μάθησης



2. Μη - Επιβλεπόμενη Μάθηση (Unsupervised Learning)

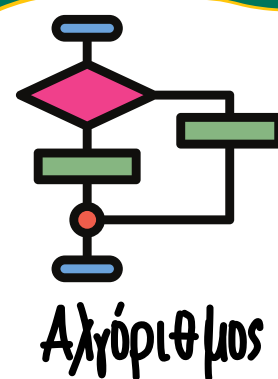
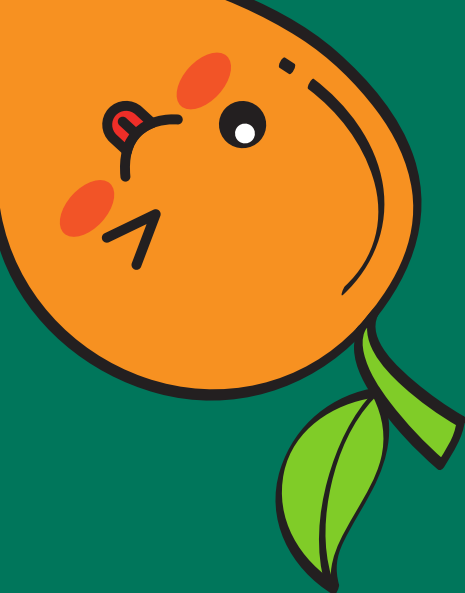


Η ανάλυση συστάδων (cluster analysis), είναι η πιο διαδεδομένη μέθοδος της μη επιβλεπόμενης μάθησης η οποία χρησιμοποιείται για την εύρεση κρυμμένων προτύπων (hidden patterns) στα δεδομένα ή να γίνει η ομαδοποίηση τους.

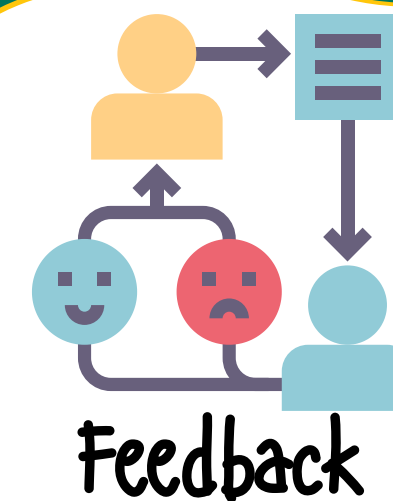


Μέθοδοι Μηχανικής Μάθησης

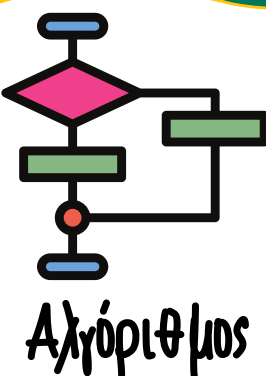
3. Ενισχυτική Μάθηση (Reinforcement Learning)



Είναι ένα Μάνγκο!



Είναι ένα Μήλο!



Όχι, είναι ένα Μήλο!



Που χρησιμοποιείται η Μηχανική Μάθηση



Βιομηχανία
Μεταφορών

Η δημιουργία μίας πιο αποτελεσματικής διαδρομής και η πρόβλεψη ενός πιθανού προβλήματος έρχεται μέσω της ανάλυσης των δεδομένων για τον εντοπισμό κρυμμένων προτύπων.



Φυσικές Πηγές
Ενέργειας

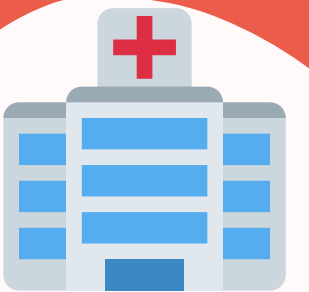
Η ανάλυση των ορυκτών του εδάφους, η διαδικασία εύρεσης για νέες πηγές ενέργειας, ακόμα και για την πρόβλεψη αστοχίας αισθητήρα διυλιστηρίου, αποτελούν κάποιες από τις πάρα πολλές εφαρμογές της μηχανικής μάθησης

Η μηχανική μάθηση αναλύει τα δεδομένα από το ιστορικό των αγορών μας και με βάση αυτά μας προτείνει παρεμφερές προϊόντα.



Λιανεμπόριο

Η ανάλυση των δεδομένων ωφελεί τον ιατρικό τομέα έτσι ώστε να γίνεται μία ταχύτερη ανίχνευση των ασθενειών, να έχουμε την δυνατότητα μιας καλύτερης θεραπείας και κυρίως την μείωση του κόστους.

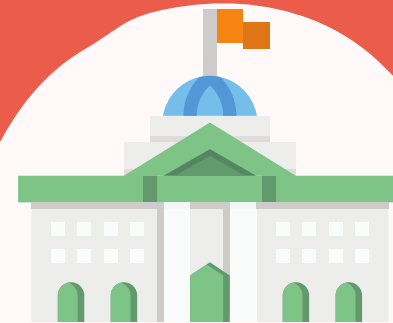


Φροντίδα
Υγείας



Χρηματοοικονομικές
Υπηρεσίες

Πρόβλεψη παράνομων ενεργειών, επενδυτικές ευκαιρίες, δημιουργία προφίλ πελατών τα οποία μπορούν να χαρακτηριστούν ως υψηλού κινδύνου.



Κυβέρνηση

Μπορεί να βοηθήσει τις κυβερνήσεις να εντοπίσουν τυχόν απάτες και να ελαχιστοποιήσουν την πιθανότητα κλοπής προσωπικών δεδομένων

Διαφορές μεταξύ της Εξαγωγής Δεδομένων, της Μηχανικής Μάθησης και της Βαθιάς Μάθησης



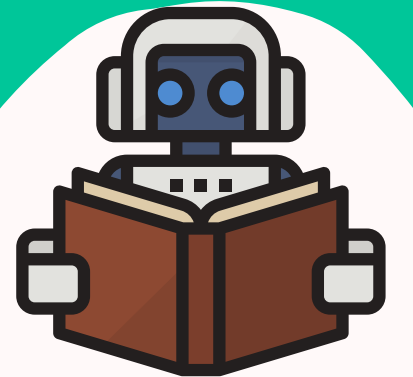
Data Mining

Η εξαγωγή δεδομένων μπορεί να θεωρηθεί ένα υπεσύνολο πολλών διαφορετικών μεθόδων για την εξαγωγή πληροφοριών από δεδομένα.



Deep Learning

Η βαθιά μάθηση συνδυάζει την πρόοδο της υπολογιστικής ισχύος και των ειδικών τύπων νευρωνικών δικτύων, ώστε να μαθαίνει περίπλοκα πρότυπα σε μεγάλες ποσότητες δεδομένων.



Machine Learning



Τελικά γιατί είναι σημαντική η Μηχανική Μάθηση

Μπορούμε να περιμένουμε περισσότερα στο μέλλον από την μηχανική μάθηση.

Καθώς τα «μεγάλα δεδομένα» (Big Data) συνεχίζουν να μεγαλώνουν, καθώς οι υπολογιστικές μηχανές γίνονται πιο ισχυρές και οικονομικά προσιτές και καθώς οι επιστήμονες των δεδομένων (Data Scientist) συνεχίζουν να αναπτύσσουν πιο ικανούς αλγόριθμους, η μηχανική μάθηση θα οδηγήσει σε μεγαλύτερη και μεγαλύτερη αποτελεσματικότητα στην προσωπική και επαγγελματική μας ζωή.

Ο αυξανόμενος

όγκος και οι πολλαπλές κατηγορίες διαθέσιμων δεδομένων, η υπολογιστική επεξεργασία που είναι φθηνότερη και ισχυρότερη, αλλά και η οικονομικά προσιτή αποθήκευση των δεδομένων, αποτελούν παράγοντες που καθιστούν τη μηχανική μάθηση πιο δημοφιλή από ποτέ.

Χάρη σε αυτήν μπορούμε να παράγουμε γρήγορα και αυτόματα μοντέλα που μπορούν να αναλύουν μεγαλύτερα και πιο περίπλοκα δεδομένα, αλλά και να παρέχουν ταχύτερα, πιο ακριβή αποτελέσματα, ακόμη και σε πολύ μεγάλη κλίμακα. Και δημιουργώντας μοντέλα ακριβείας, ένας οργανισμός έχει πολύ καλύτερες πιθανότητες να εντοπίζει επικερδείς ευκαιρίες, ή να αποφεύγει άγνωστους κινδύνους.

