

# Εξόρυξη δεδομένων με χρήση τεχνικών μηχανικής μάθησης

Τμήμα Μηχανικών Πληροφορικής και Υπολογιστών



# Εισαγωγή

# Για ποιο λόγο να παρακολουθήσω....

- Επειδή θέλω να γίνω αναλυτής δεδομένων



- Επειδή στην εταιρία που εργάζομαι θα πρέπει να παίρνω γρήγορες και σωστές αποφάσεις

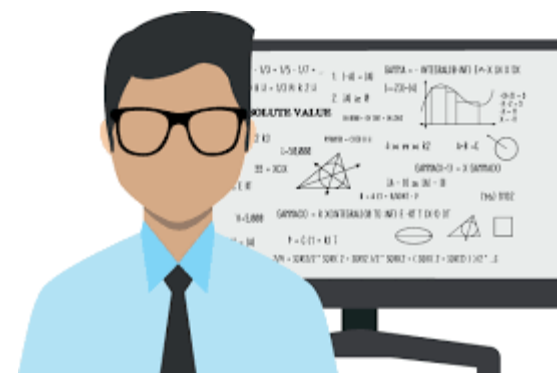


# Για ποιο λόγο να παρακολουθήσω....

- Επειδή μου αρέσει να ανακαλύπτω νέα αντικείμενα.



- Γιατί είμαι αναλυτής δεδομένων και θέλω να επικαιροποιήσω τις γνώσεις μου



# Τι χρειαζομαι για να ξεκινήσω

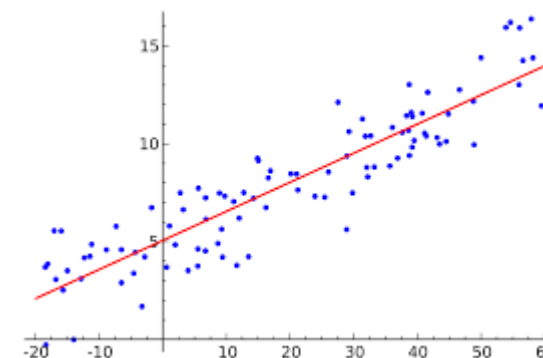
- Αναλυτική Σκέψη



- Γνώσεις Η/Υ



- Βασικές Γνώσεις Μαθηματικών (επίπεδο Λυκείου)



# Τι δεν είναι απαραίτητο....

- Να είμαι προγραμματιστής



- Να γνωρίζω καλά μαθηματικά

A chalkboard filled with mathematical formulas. The most prominent formula is the derivative of the log-likelihood function for a normal distribution:  $\frac{\partial}{\partial a} \ln f_{a, \sigma^2}(\xi_1) = \frac{(\xi_1 - a)}{\sigma^2} f_{a, \sigma^2}(\xi_1) - \frac{1}{2\sigma^2} \left( \frac{\xi_1 - a}{\sigma^2} \right)^2$ . Other formulas include the expectation of the score function:  $\int T(x) \cdot \frac{\partial}{\partial \theta} f(x, \theta) dx = M \left( T(\xi) \cdot \frac{\partial}{\partial \theta} \ln L(\xi, \theta) \right)$  and the Fisher information:  $\int T(x) \cdot \left( \frac{\partial}{\partial \theta} \ln L(x, \theta) \right) \cdot f(x, \theta) dx = \int T(x) \cdot \left( \frac{\partial}{\partial \theta} f(x, \theta) \right) dx$ . The word "Life" is written in a simple font at the bottom left of the chalkboard.

- Ολοκληρώνοντας το Πρόγραμμα Θα Μπορείς να.....

# Να Σχεδιάζεις μοντέλα ανάλυσης





# Να διαβάξεις δεδομένα και να αποφασίζεις



# Πιο συγκεκριμένα

- Έξυπνη χρήση των Κοινωνικών Δικτύων (π.χ., στοχευμένη προώθηση προϊόντων)



- Χρηματοοικονομική Ανάλυση



# Πιο συγκεκριμένα

- Επιλογή του ιδανικού υποψηφίου



- Σύσταση προϊόντων



# Ενότητες



# Τρόπος Διεξαγωγής

- Μετά το τέλος κάθε διδακτικής ενότητας θα υπάρχουν τεστ αξιολόγησης που θα αφορούν θέματα εμβάθυνσης και κατανόησης των διδασκόμενων εννοιών.
- Τα τεστ θα υποβάλλονται στην ηλεκτρονική πλατφόρμα του προγράμματος.
- Στο τέλος κάθε τεστ θα βλέπετε και τους βαθμούς που συγκεντρώσατε.

# Εργαλείο

- RapidMiner Community Edition



# Μαθήματα για Data Scientist



Data Mining &  
Machine Learning



Social Network  
Analysis



R και Python



Σχεδιασμός και  
Διαχείριση  
Data-Warehouse  
με PL/SQL



Διαχείριση μεγάλων  
Δεδομένων  
(Big Data) με  
Hadoop/Spark

# Σας ευχαριστώ

A. Τσολακίδης, PhD

[atsolakid@uniwa.gr](mailto:atsolakid@uniwa.gr)

[datascience@uniwa.gr](mailto:datascience@uniwa.gr)

<http://datascience.uniwa.gr/>

