COMPUTER SCIENCE

UNIT I WEEK 4, TUESDAY MAY 8TH + THURSDAY MAY 10TH

WWW.CITA.UTORONTO.CA/~WOODFORD

WOODFORD@CITA.UTORONTO.CA

THIS WEEK IN CS AND STEM

- Hawking's final theory
  - https://futurism.com/stephen-hawkings-final-theory-multiverse/
- A CRISPR app
  - https://futurism.com/mammoth-biosciences-crispr-diagnostics/
- Military AI learning to write software
  - https://futurism.com/military-created-ai-learned-to-program/
- They're not AI scooters, but still a very interesting tech issue
  - https://futurism.com/scooter-not-future-transportation/

FINAL PROJECTS

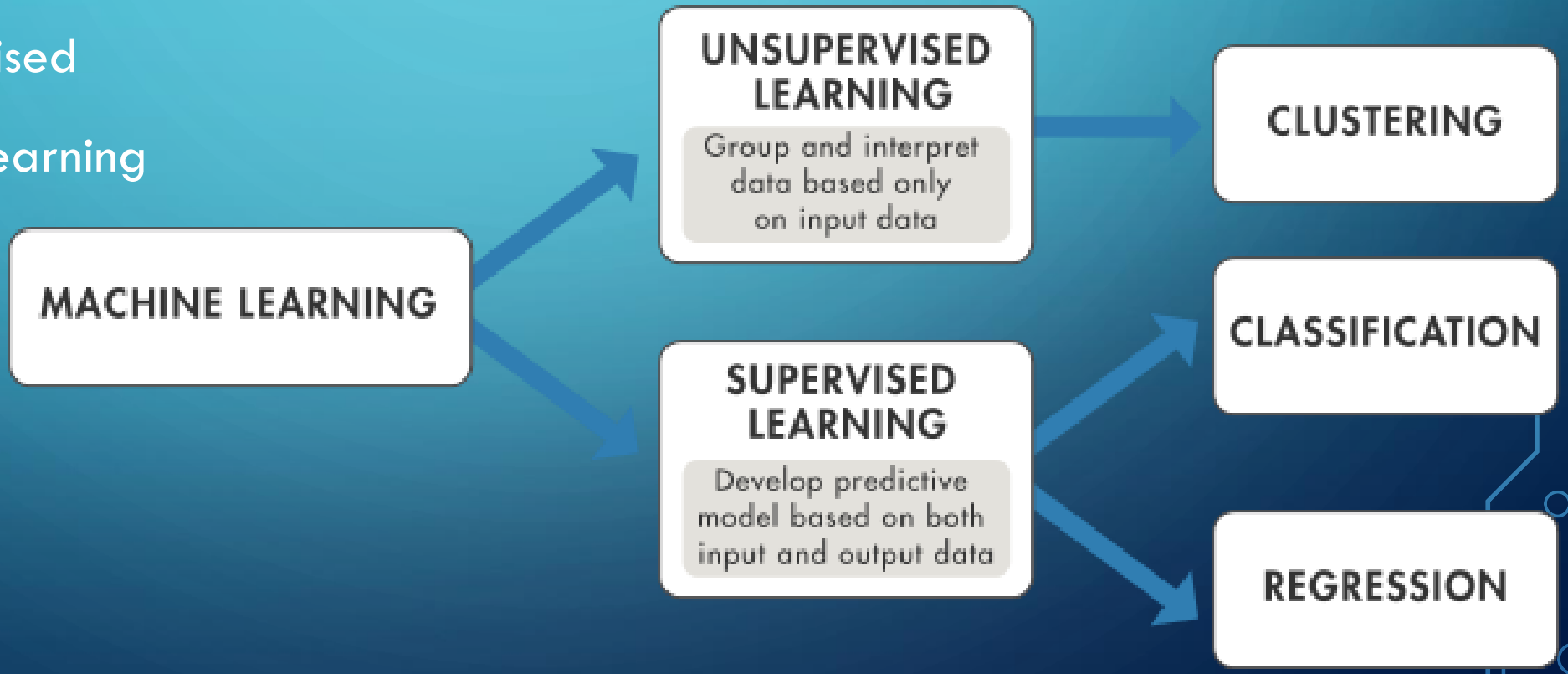
- 3 different project descriptions, you need to:
  - Complete the coding, using all of the skills we've learned
  - Write user documentation
  - Write a report
  - Give a 10 minute presentation showcasing your work (last day of classes, June 5th)
- Due dates:
  - ~~Pick your project (1 of the 3 given): April 3rd – tell me in person or in email!~~
  - **Update your github account regularly as you add sources, make changes, and code.**
  - Submit Report, code, supporting documents: June 4th at midnight
  - Give presentation: June 5th in class

STUDENT-LED LESSONS

- The last 2-3 classes will be based on what YOU want to learn about relating to Computer Science.
- Aim for 1/3 -½ class per student, maximum of 2 topics per student.
- Send your topics to me by May 15th . Looking at May 24, 29 for topic discussions.

TYPES OF MACHINE LEARNING

- Supervised
- Unsupervised
- Semi-supervised
- Reinforced learning



SUPERVISED MACHINE LEARNING TECHINQUES

- Classification: support vector machine (SVM), boosted and bagged decision trees, k-nearest neighbour, naïve bayes, discriminant analysis, logistic regression, neural networks.

- Regression: linear model, nonlinear model, regularization, stepwise regression, boosted and bagged decision trees, neural networks, adaptive neuro-fuzzy learning.

UNSUPERVISED MACHINE LEARNING TECHNIQUES

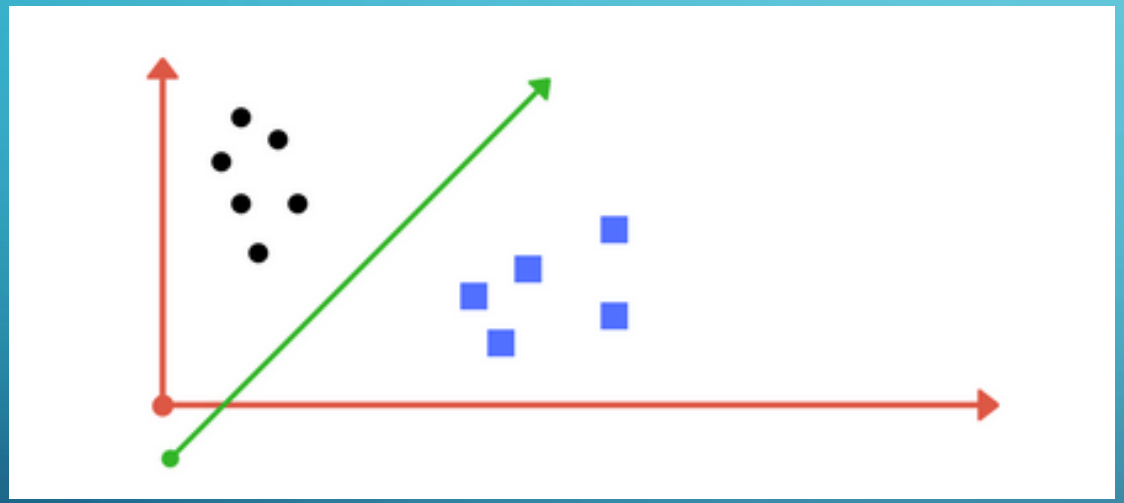
- Clustering: k-means and k-mediods, hierarchical clustering, Gaussian misture models, hidden Markov models, self-organizing maps, fuzzy c-means clustering, subtractive clustering

POPULAR METHODS

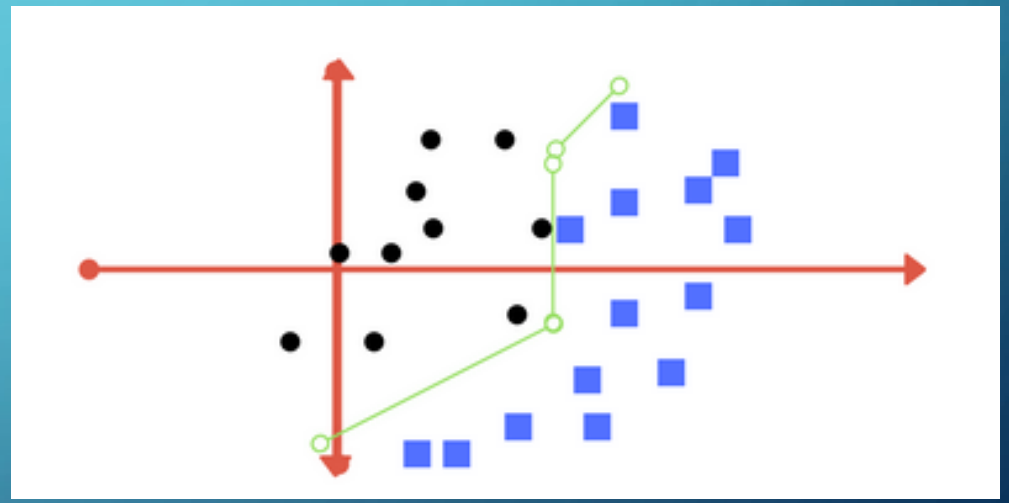
- Support Vector Machines (SVM)
- Bagged and Boosted Decision Trees
- K-nearest neighbour
- Linear Regression
- Logistic Regression
- Naïve Bayes
- Linear Discriminant Analysis
- K-means
- Dimensionality reduction
- Neural Networks

SUPPORT VECTOR MACHINES (SVM)

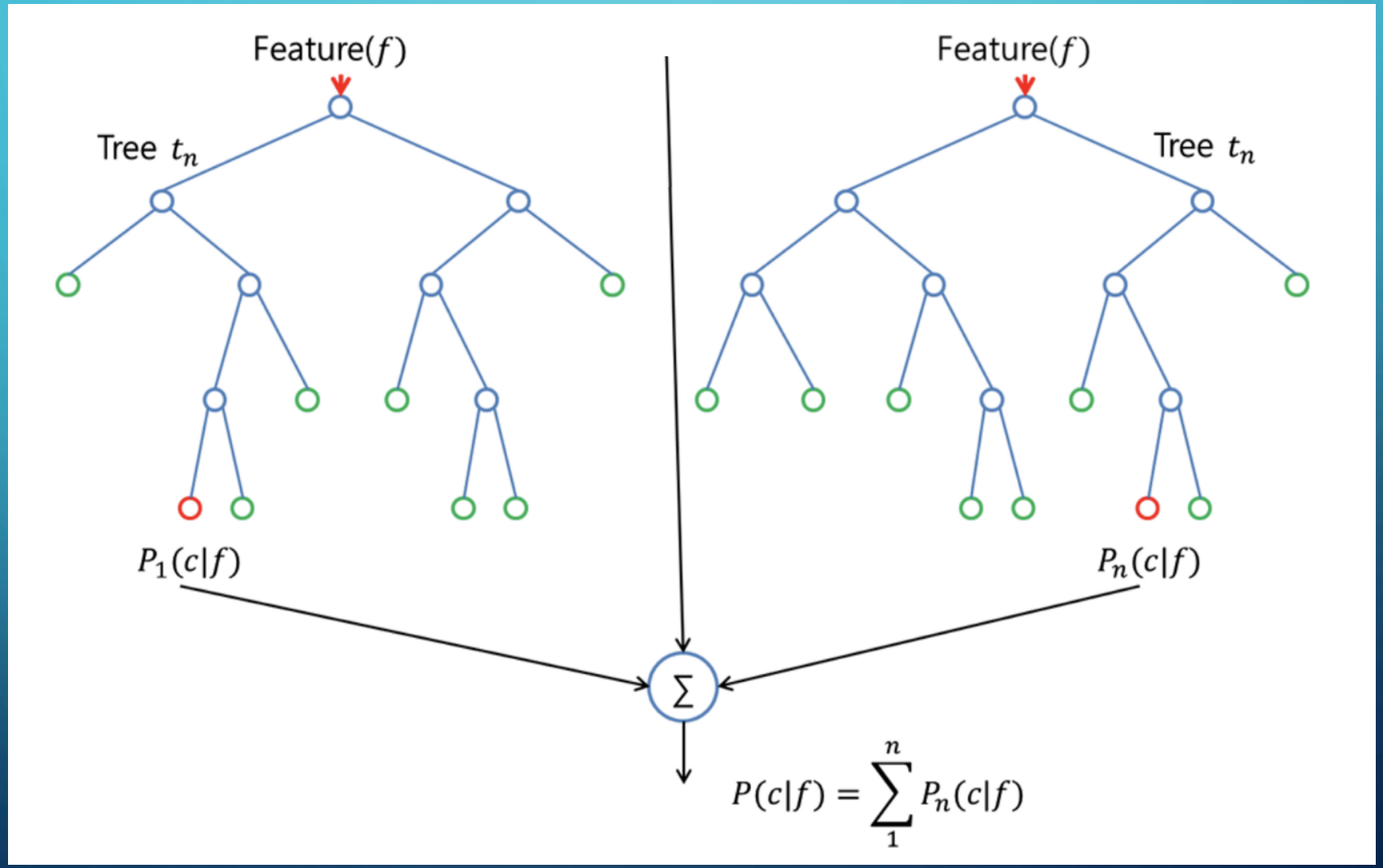
- Simple Binary Case



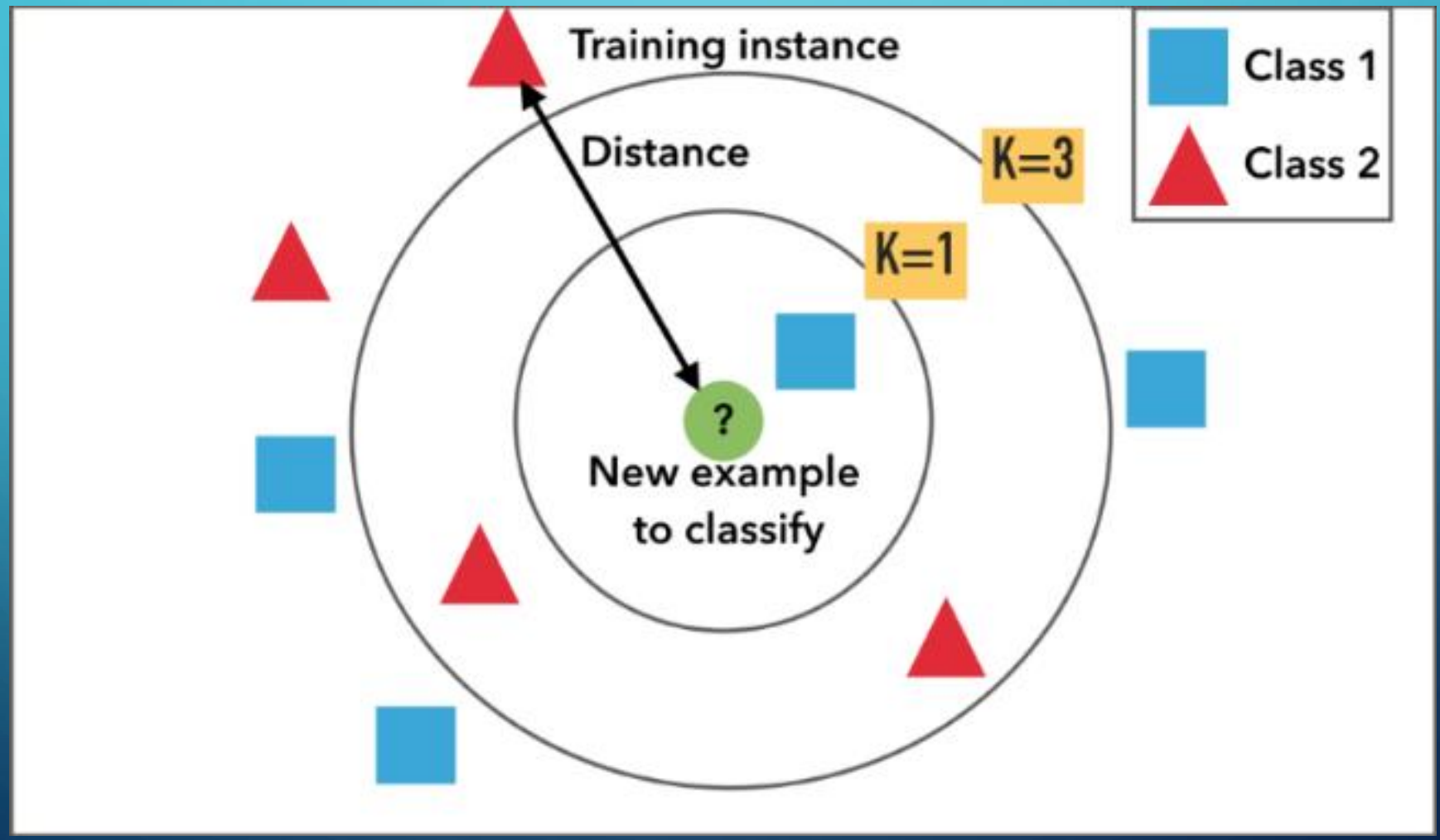
- Nonlinear Binary Case



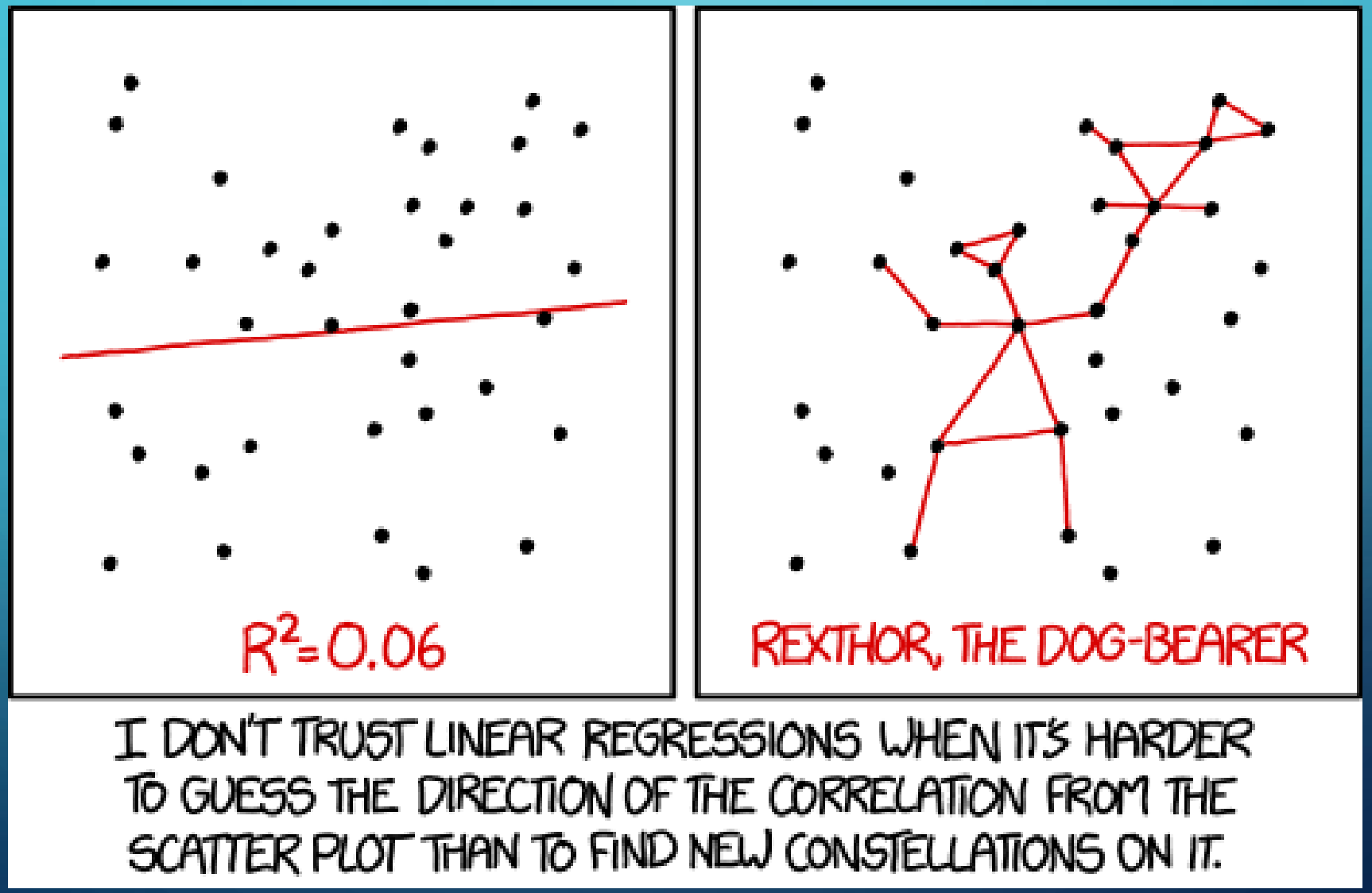
BOOSTED AND BAGGED DECISION TREES



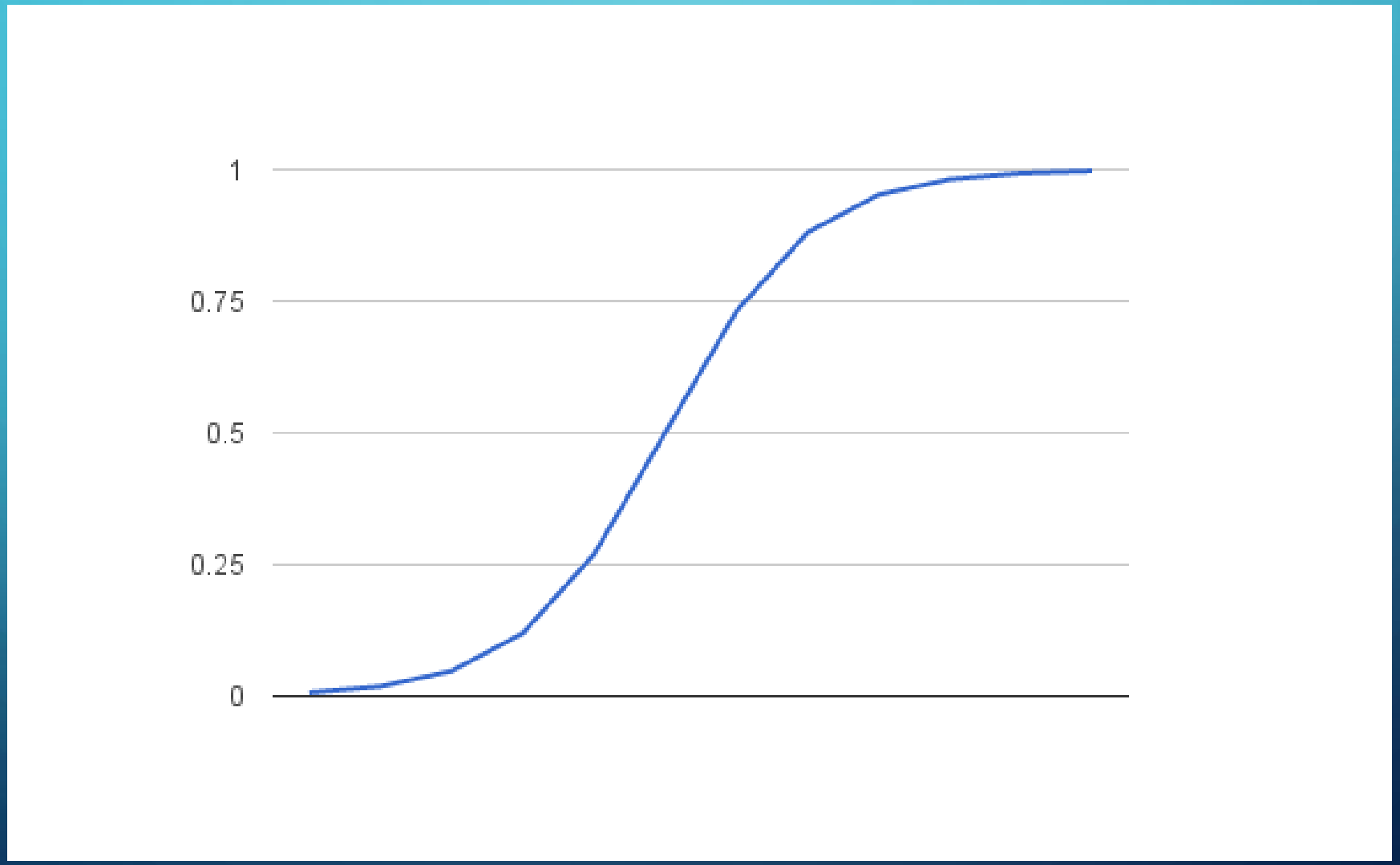
K-NEAREST NEIGHBOUR



LINEAR REGRESSION



LOGISTIC REGRESSION



REFERENCES AND RESOURCES

- https://towardsdatascience.com/supervised-vs-unsupervised-learning-14f68e32ea8d
- https://www.mathworks.com/discovery/machine-learning.html
- https://www.toptal.com/machine-learning/machine-learning-theory-an-introductory-primer
- https://www.youtube.com/watch?v=UqYde-LULfs
- https://towardsdatascience.com/a-tour-of-the-top-10-algorithms-for-machine-learning-newbies-dde4edffae11
- https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/
- https://www.youtube.com/watch?v=1NxnPkZM9bc
- https://medium.com/machine-learning-101/chapter-2-svm-support-vector-machine-theory-f0812effc72
- https://towardsdatascience.com/the-random-forest-algorithm-d457d499ffcd
- https://developers.google.com/machine-learning/crash-course/prereqs-and-prework
- https://machinelearningmastery.com/logistic-regression-for-machine-learning/