MESS 2019

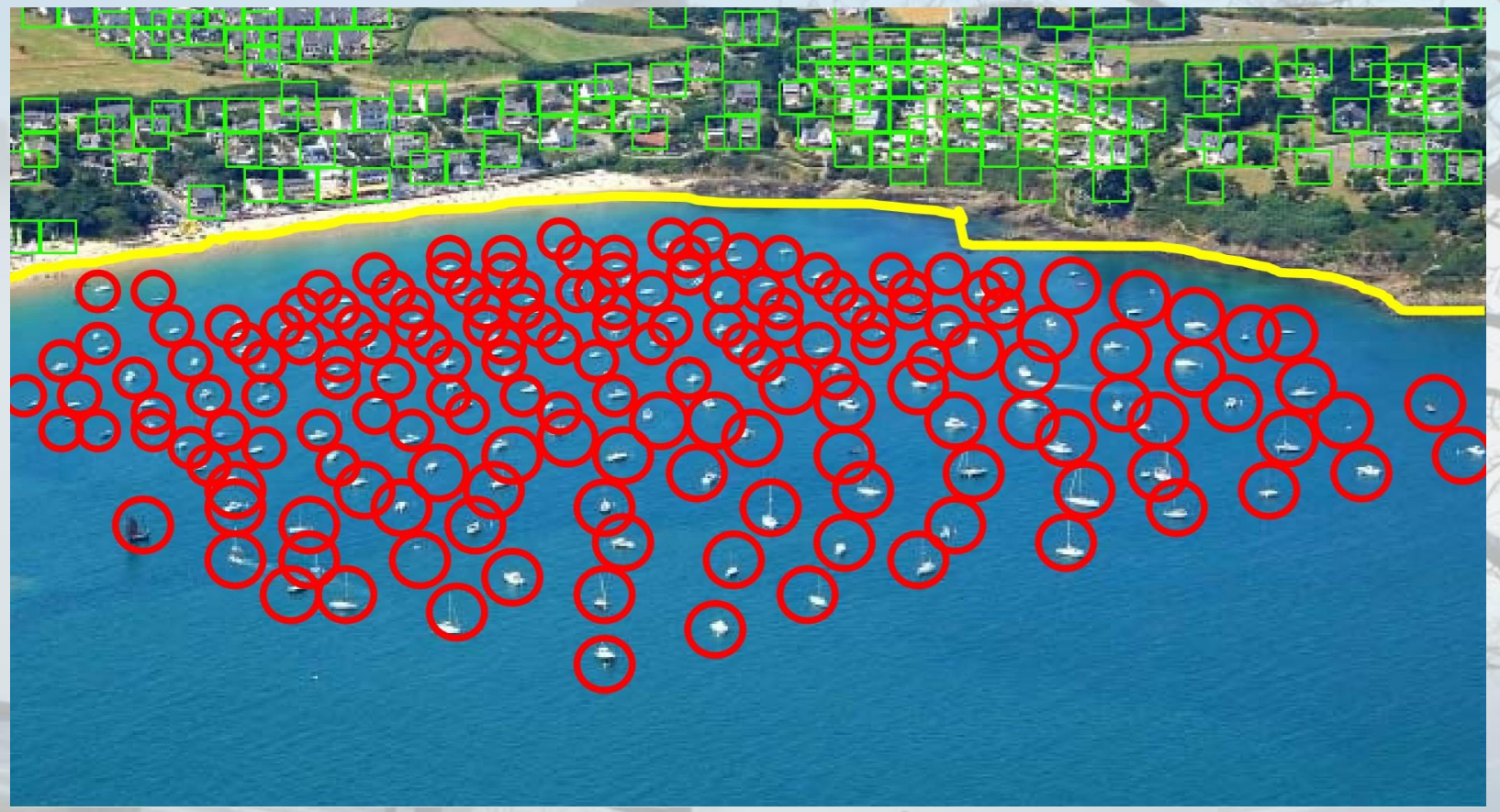
Support Vector Machines

Polina Polunina

Basic Principles of Classification

- Want to classify objects as boats and houses

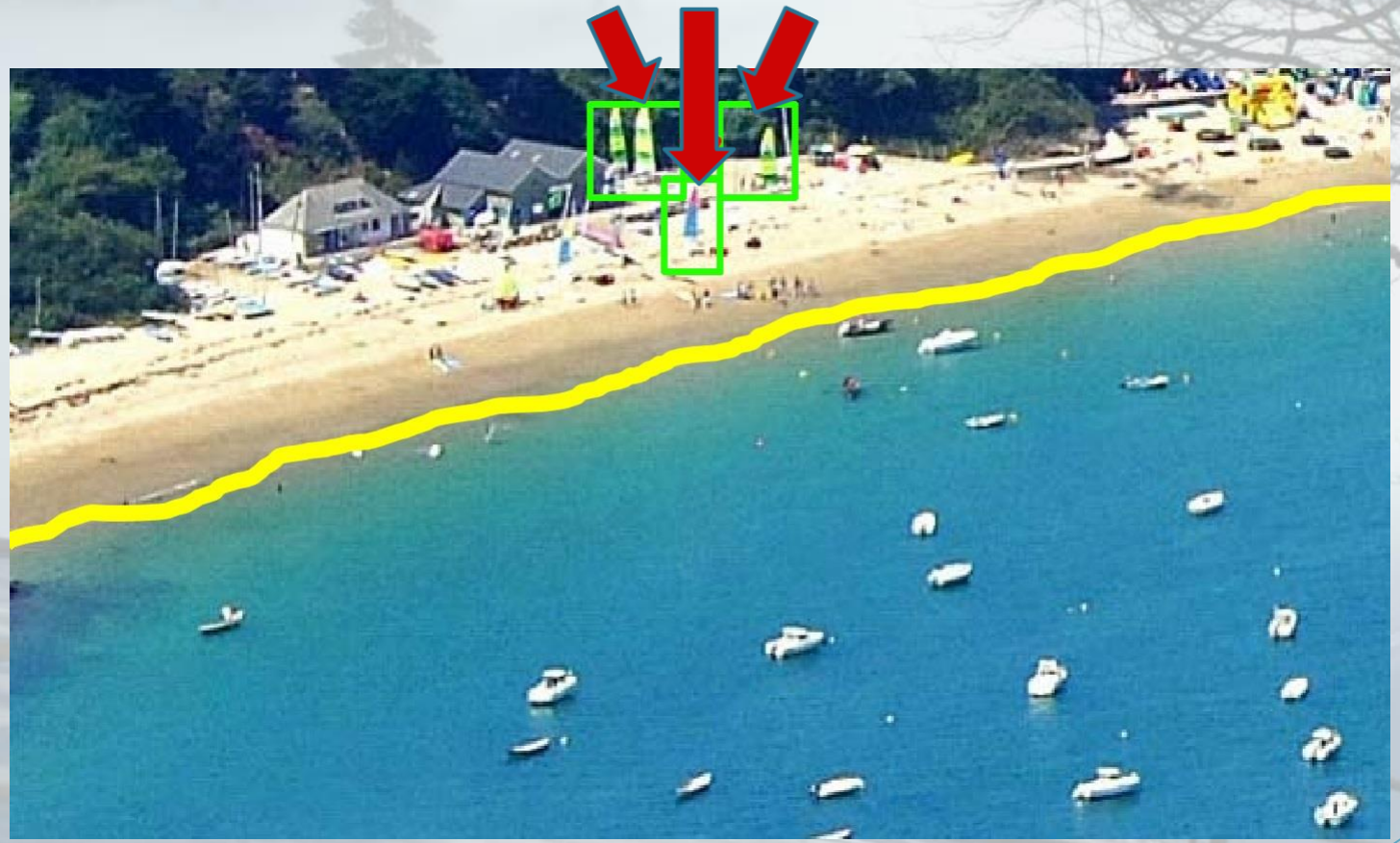
Basic Principles of Classification



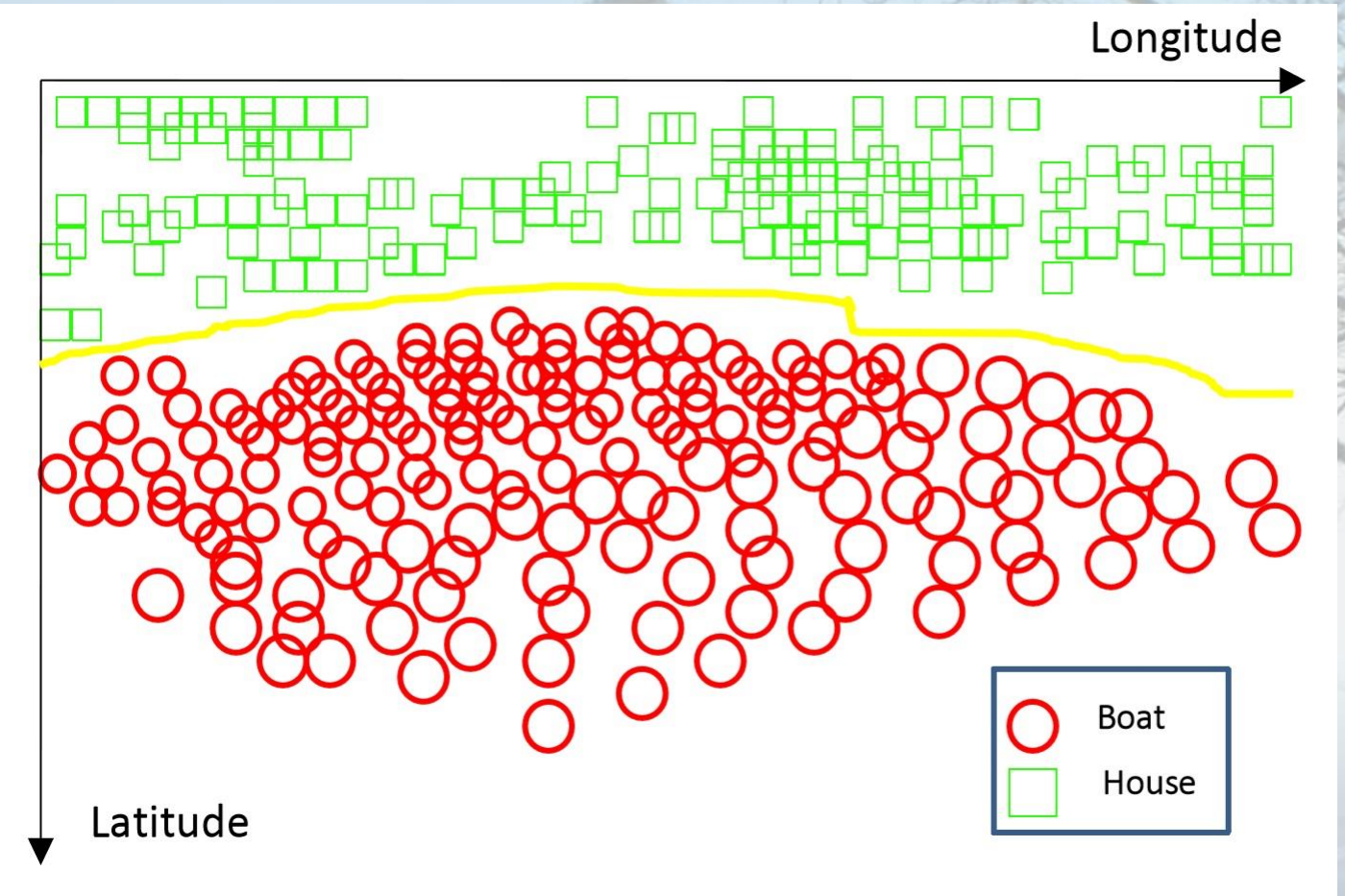
- All objects before the coast line are boats and all objects after the coast line are houses
- Coast line serves as a decision surface that separates two classes

Basic Principles of Classification

- These boats will be misclassified as houses

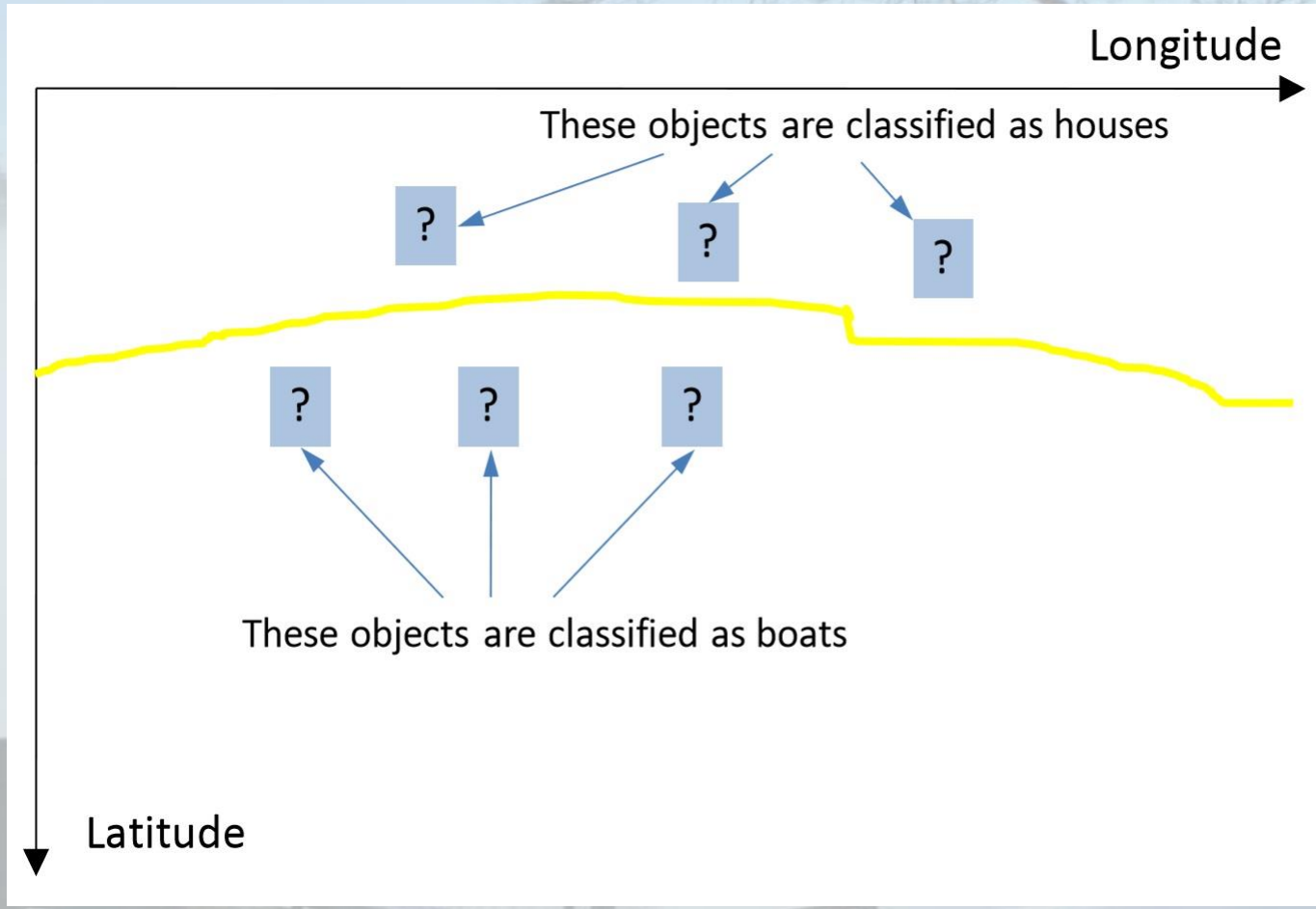


Basic Principles of Classification



- First all objects are represented mathematically
- Then the algorithm seeks to find a decision surface that separates classes of objects

Basic Principles of Classification



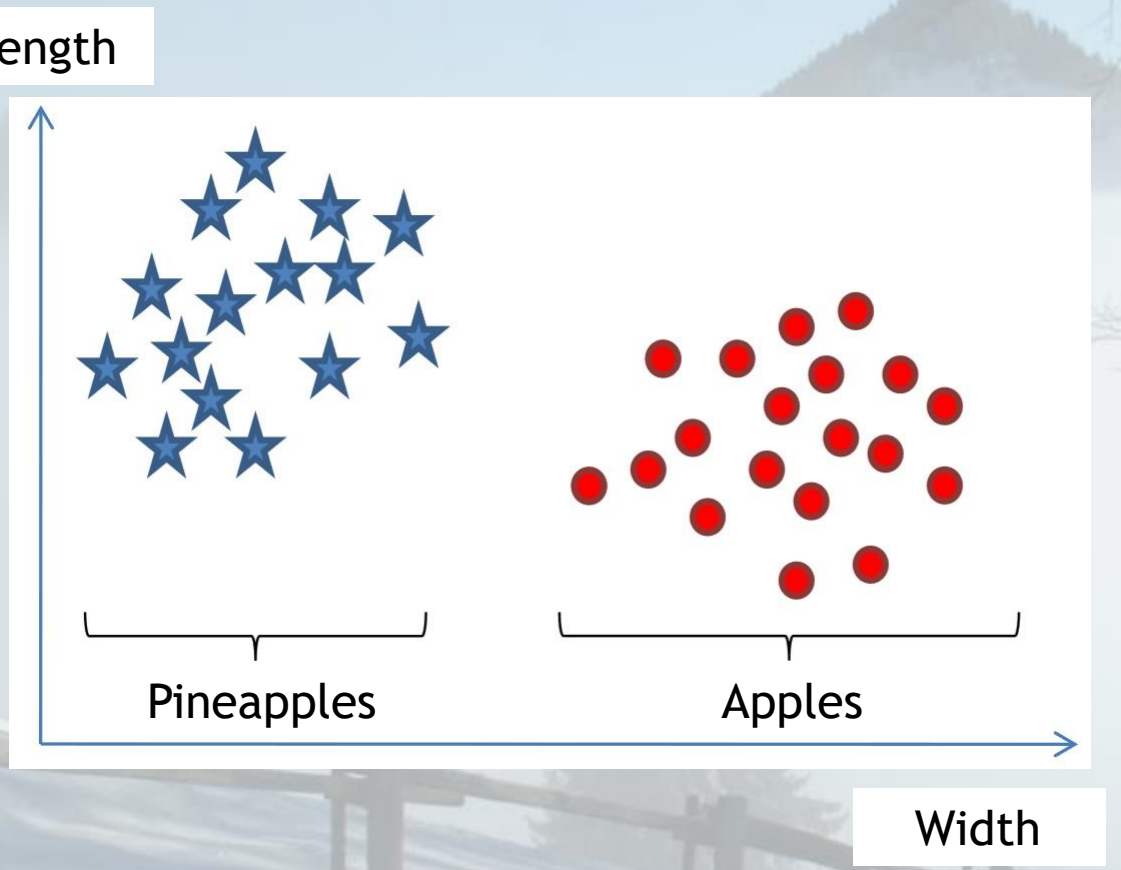
- **New (previously unseen) objects that are below the decision surface will be classified as “boats” and the objects above the surface will be classified as “houses”**

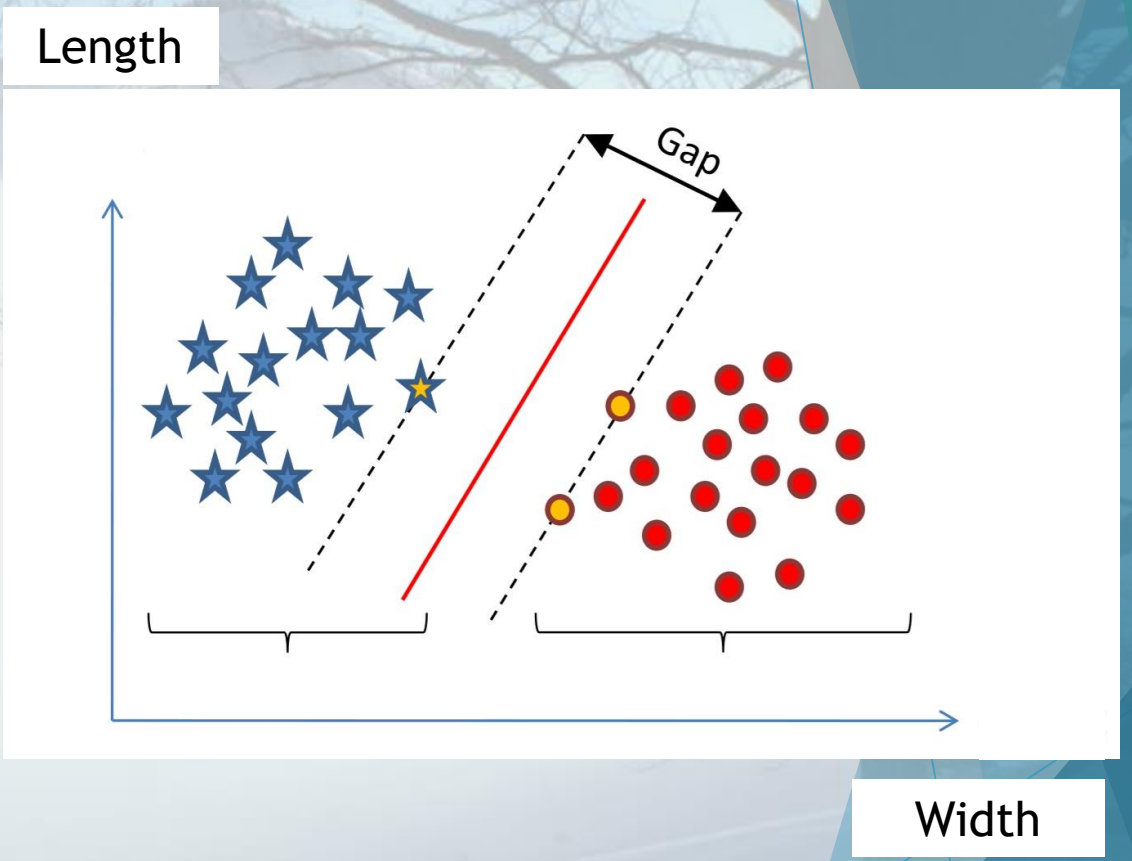
Support Vector Machines

- SVM is a ML algorithm that offers a solution to regression and classification problems

- Importance of SVM:
  - Robust to very large number of variables and small samples
  - Can learn both simple and highly complex models
  - Employ sophisticated mathematical principles to avoid overfitting
  - Provides superior empirical results

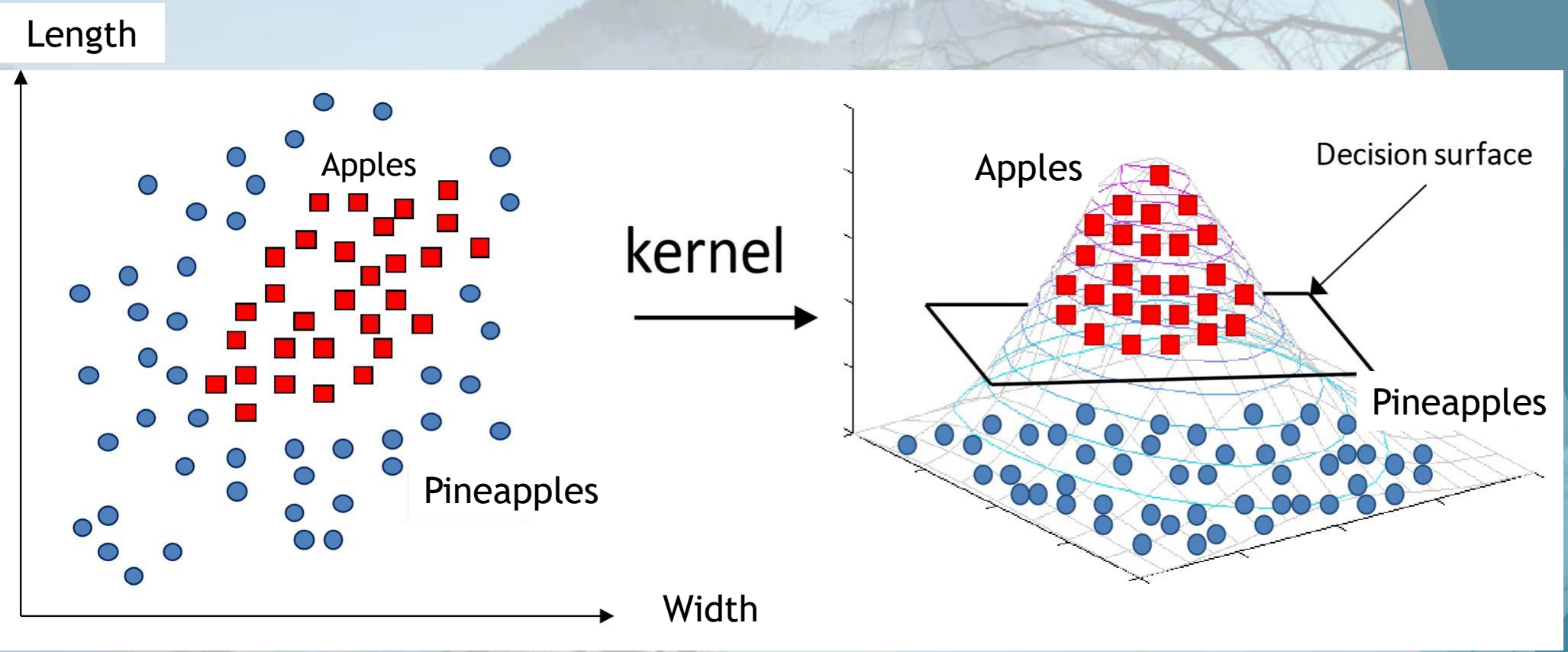
Support Vector Machines





Goal: Find a hyperplane that **can separate** the classes **AND** has **the largest gap** (a.k.a. margin maximization) between borderline data points (a.k.a. support vectors)

Support Vector Machines



If the data is not linearly separable, we can use nonlinear mappings into a higher dimensional space (a.k.a. kernels) where the data become separable