Support Vector Machine (SVM)
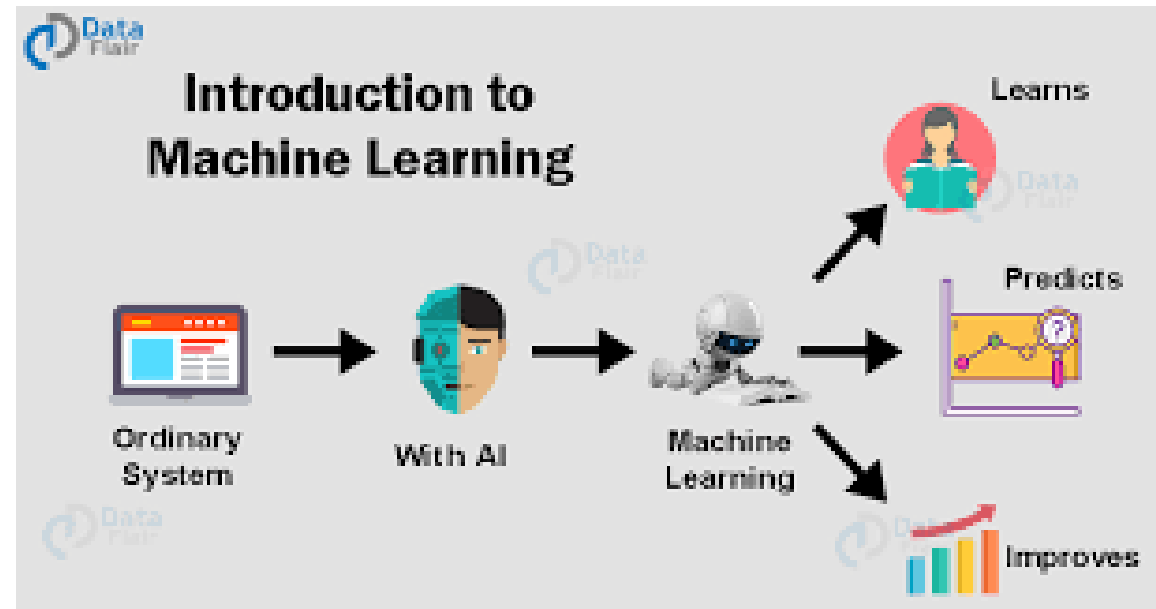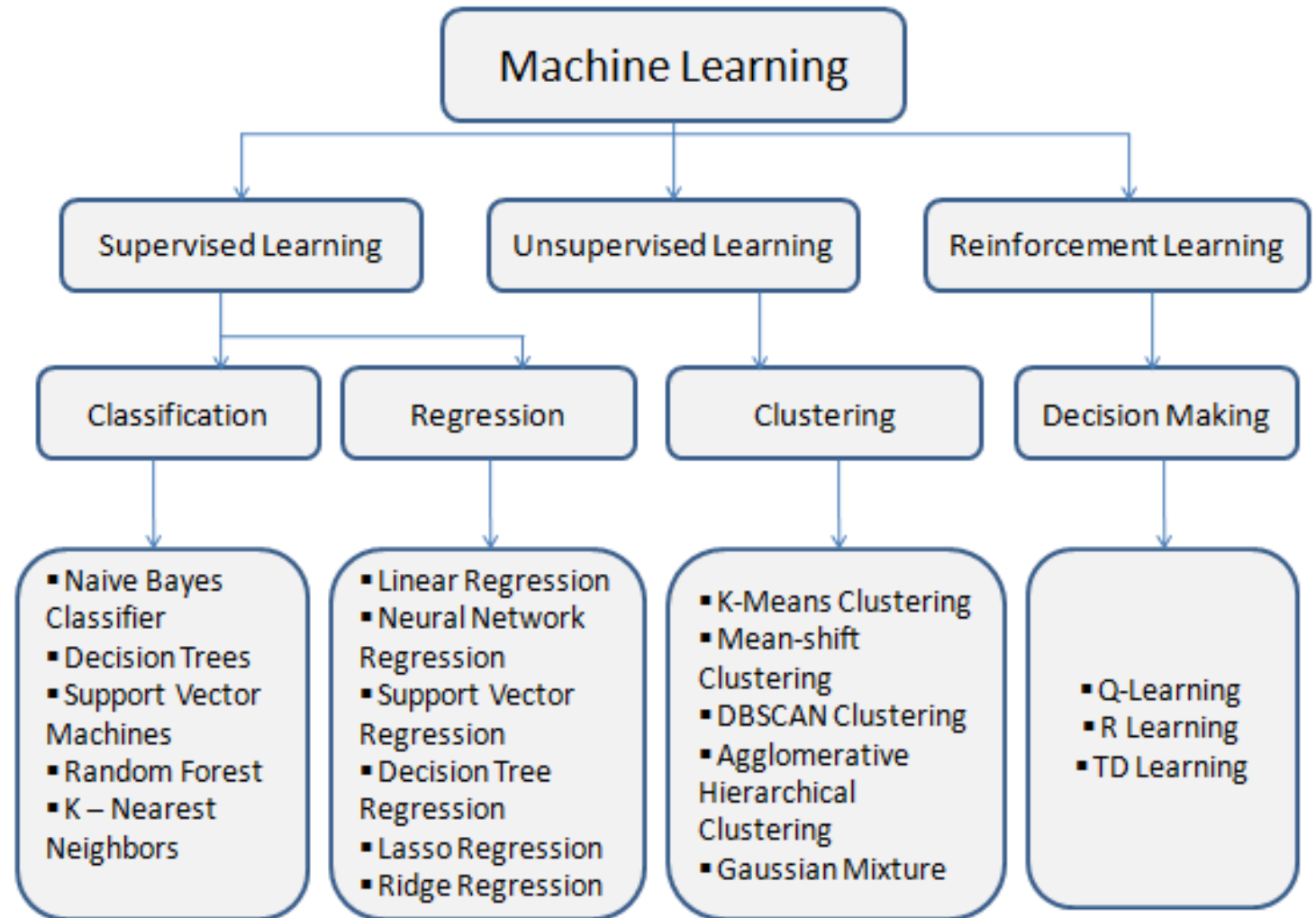
Machine Learning
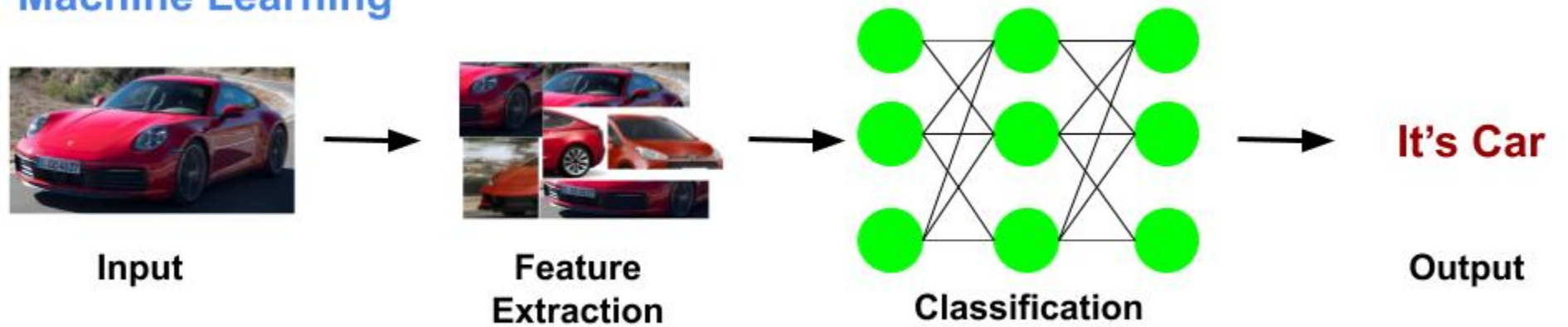
Types of Machine Learning Algorithms

Machine Learning

- **Supervised Learning**
  - **Classification**
    - Naive Bayes Classifier
    - Decision Trees
    - Support Vector Machines
    - Random Forest
    - K – Nearest Neighbors
  - **Regression**
    - Linear Regression
    - Neural Network Regression
    - Support Vector Regression
    - Decision Tree Regression
    - Lasso Regression
    - Ridge Regression
- **Unsupervised Learning**
  - **Clustering**
    - K-Means Clustering
    - Mean-shift Clustering
    - DBSCAN Clustering
    - Agglomerative Hierarchical Clustering
    - Gaussian Mixture
- **Reinforcement Learning**
  - **Decision Making**
    - Q-Learning
    - R Learning
   - TD Learning

Agenda

- 1. Support Vectors
- 2. Hyperplanes
- 3. Marginal Distance
- 4. Linear Separable
- 5. Non-Linear Separable

## Support Vector Machine Algorithm (Supervised Learning - Classification)

- Support Vector Machine algorithms are supervised learning models that analyze data used for classification and regression analysis.

- They essentially filter data into categories, which is achieved by providing a set of training examples, each set marked as belonging to one or the other of the two categories.

- The algorithm then works to build a model that assigns new values to one category or the other.
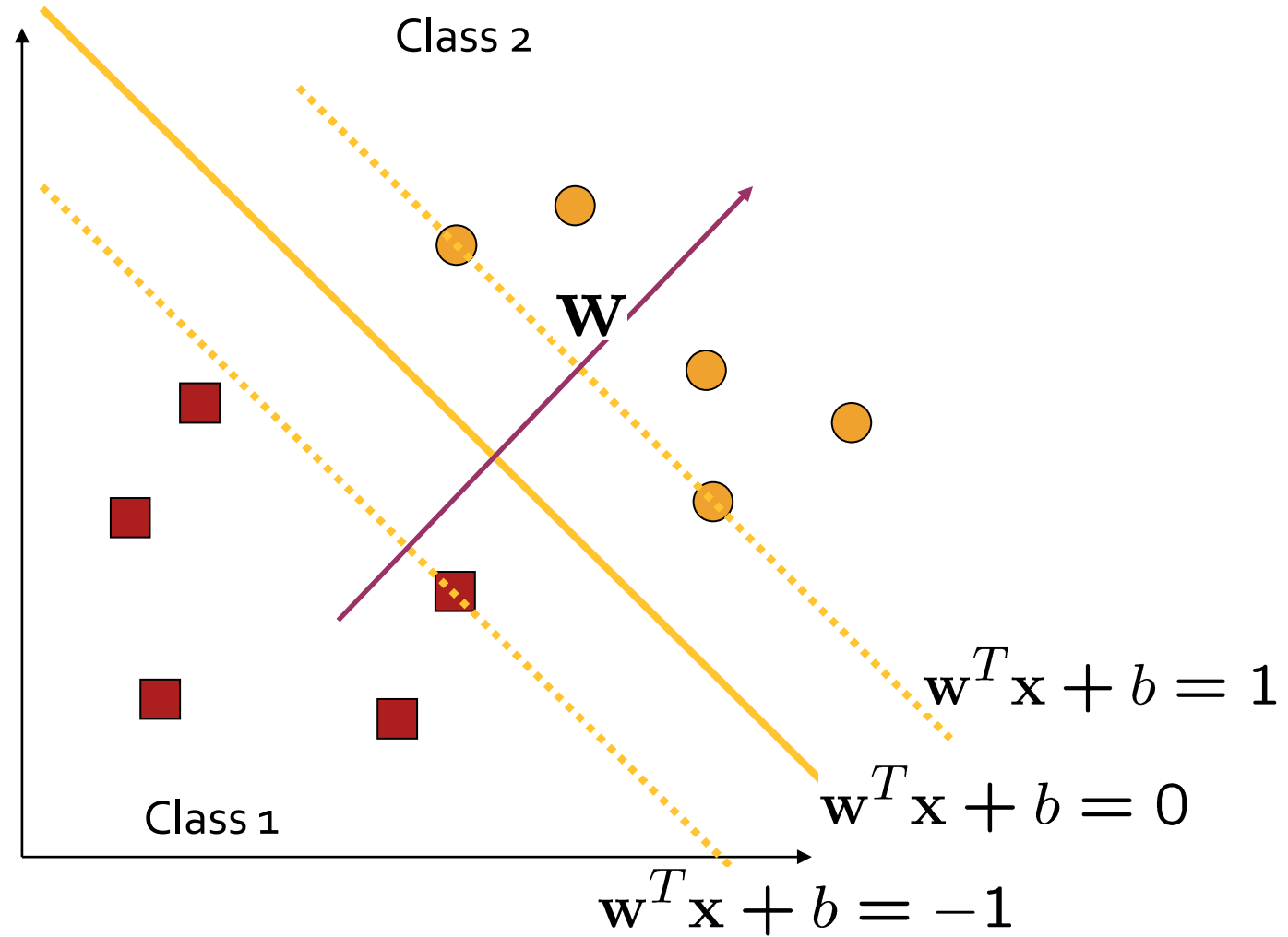
Linear Separable Points
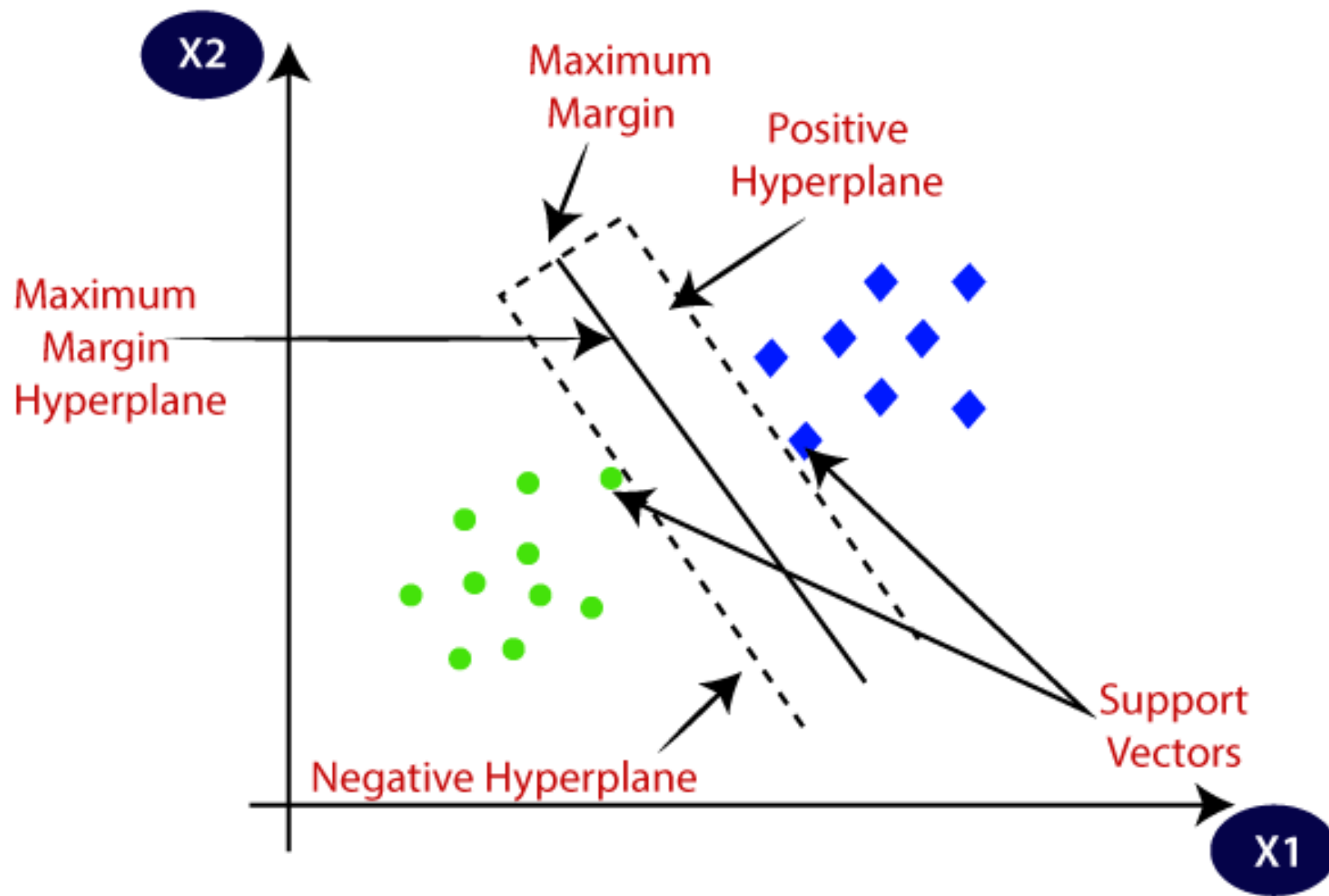
Non-Linear Separable Points
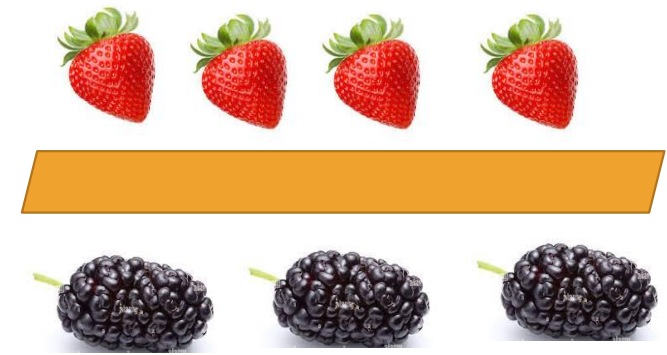
A Geometrical Interpretation

Class 2

Class 1

$\mathbf{W}$

$\mathbf{w}^T\mathbf{x} + b = 1$

$\mathbf{w}^T\mathbf{x} + b = 0$

$\mathbf{w}^T\mathbf{x} + b = -1$

Software

7 Free SVM TOOLS

- KNIME

- LIBSVM

- mySVM

- Orange

- RapidMiner

- SVM Light

- Weka

Steps for Classification

- Prepare the pattern matrix

- Select the kernel function to use

- Select the parameter of the kernel function and the value of C

- You can use the values suggested by the SVM software, or you can set apart a validation set to determine the values of the parameter

- Execute the training algorithm and obtain the α_i

- Unseen data can be classified using the α_i and the support vectors

Advantages

- Effective in high dimensional spaces.

- Still effective in cases where number of dimensions is greater than the number of samples.

- Uses a subset of training points in the decision function (called support vectors), so it is also memory efficient.

- Versatile: different Kernel functions can be specified for the decision function. Common kernels are provided, but it is also possible to specify custom kernels.

Disadvantages

- If the number of features is much greater than the number of samples, avoid over-fitting in choosing Kernel functions and regularization term is crucial.

- SVMs do not directly provide probability estimates, these are calculated using an expensive five-fold cross-validation

Conclusion

- SVM is a useful alternative to neural networks

- Two key concepts of SVM: maximize the margin and the kernel trick

- Many SVM implementations are available on the web for us to try on our data set!

References

- https://scikit-learn.org/stable/modules/svm.html#complexity

- http://www.butleranalytics.com/7-free-svm-tools/

- https://www.sas.com/en_gb/insights/articles/analytics/machine-learning-algorithms.html

- https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/

- https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/