A geometrical approach to ML

1. Draw a line that sits as far as possible from the data points $\rightarrow$ Support Vector Machines
2. Send all data points in a higher dimension space where they are linearly separable $\rightarrow$ kernel trick

$\Rightarrow$ SVM + kernel trick = Find the optimal separating hyperplane in this higher dimension space, without ever computing the mapping.

- SVM try to separate data by maximizing a geometrical margin
- They are computed offline
- They offer a sparse, robust to class imbalance, and easy to evaluate predictor
- Kernels are a way of enriching (lifting) the data representation so that it becomes linearly separable
- SVMs + kernels offer a versatile method for classification, regression and density estimation
- [Link] to documentation in scikit-learn