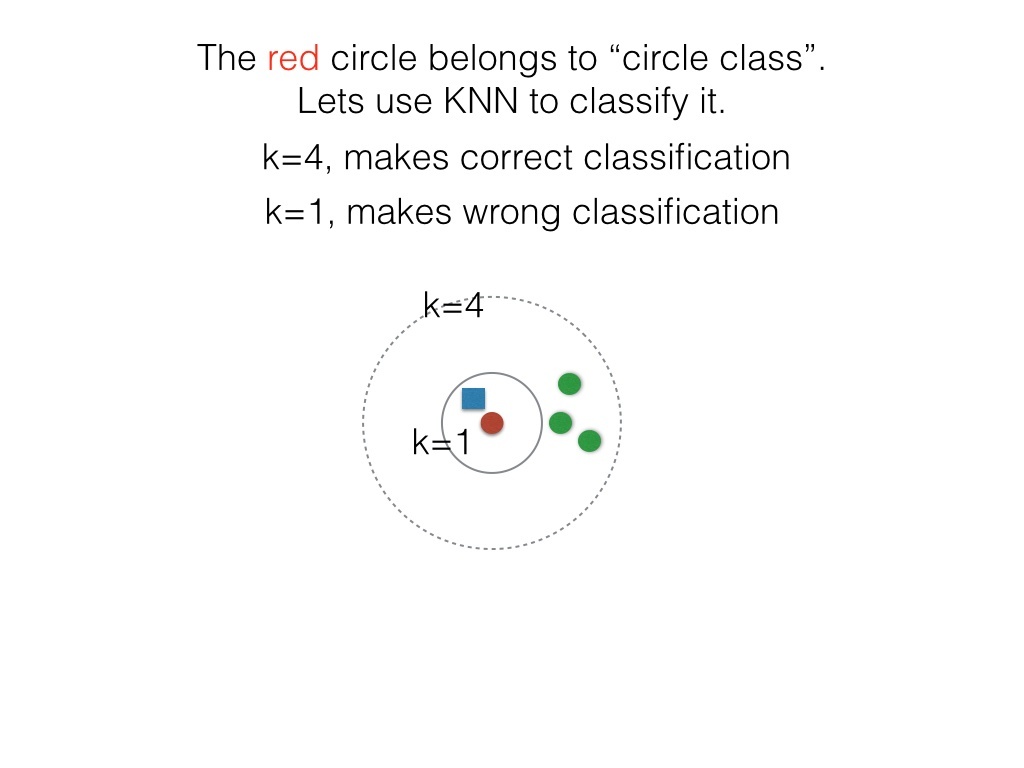
**Impact of K value in KNN algorithm on bias and variance**

* large K = siimple model = underfit = low variance & high bias
* small K = complex model =overfit = high variance & low bias
* **When K increases to inf, the model is simplest.** All test data point will belong to the same class: the majority class. This is **under-fit, that is, high bias and low variance.**
* When K decreases, **lets say K = 1 the granularity or resolution is too fine, which is overfit.**

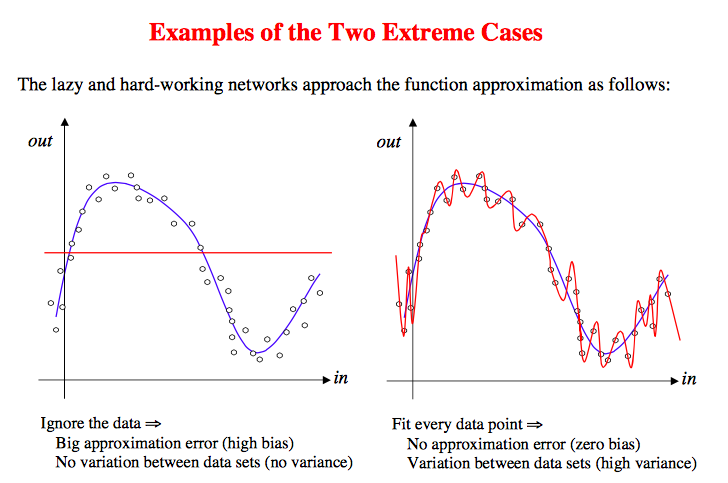
**Overfit = high variance**

Example :

The red circle as test data, is wrongly classified because the model tries too hard and its “granularity” is too fine, as a result the **“outliers” and “noised”** affect the decision process.



The following picture shows : simply averaging **all** data (**similar to a very large**KK**in KNN**) gets high bias and no variance. (Refer to [http://www.cs.bham.ac.uk/~jxb/IN...](http://www.cs.bham.ac.uk/~jxb/INC/l9.pdf), page 13)



Also note:

[Does k-NN with k=1 always implies overfitting?](http://stats.stackexchange.com/questions/107870/does-k-nn-with-k-1-always-implies-overfitting)

[k-nearest neighbors algorithm - Wikipedia](https://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm#The_1-nearest_neighbour_classifier)

“The most intuitive nearest neighbour type classifier is the one nearest neighbour classifier that assigns a point x to the class of its closest neighbour in the feature space.

As the size of training data set approaches infinity, the one nearest neighbour classifier guarantees an error rate of no worse than twice the Bayes error rate (the minimum achievable error rate given the distribution of the data).”