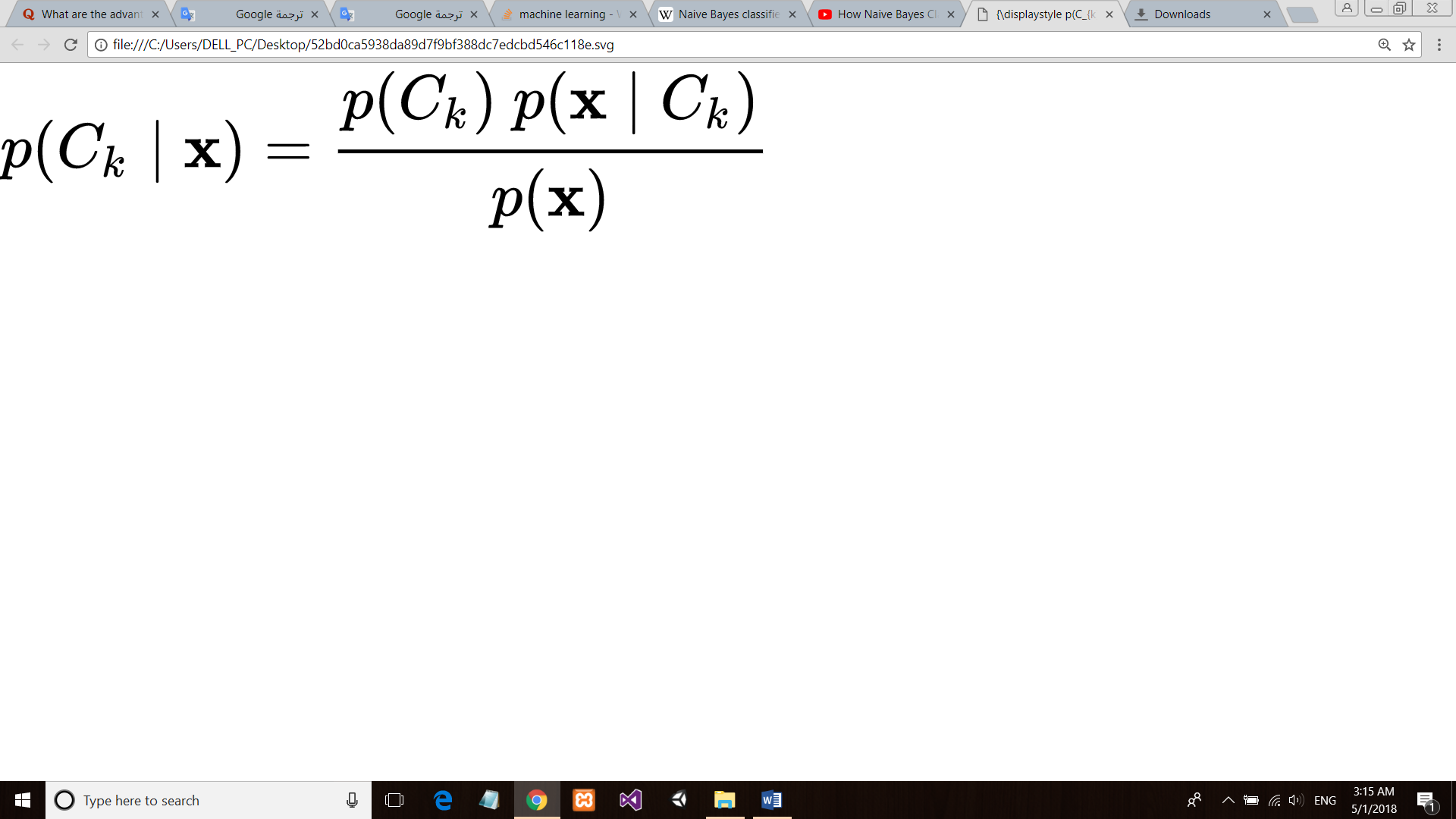
**Data set information :**

**BBC Sport data set.**

* Consists of 737 documents from the [BBC Sport](http://news.bbc.co.uk/sport1/hi/default.stm) website corresponding to sports news articles in five topical areas from 2004-2005.
* Class Labels: 5 (athletics, cricket, football, rugby, tennis)

**naive Bayes classifier:**

**rule :**



with accuracy of more than 59%

assume that the value of a particular feature is [independent](https://en.wikipedia.org/wiki/Independence_(probability_theory)) of the value of any other feature considers each of these features to contribute independently to the probability , regardless of any possible [correlations](https://en.wikipedia.org/wiki/Correlation_and_dependence) between those features

as it has no complicated iterative parameter estimation which made it easy to build and it also can outperform more classifications methods.

**Support victor machine classifier:**

with accuracy of more than 71%

as it works as a data point is viewed as a P{\displaystyle p}PP-dimensional vector (a list of P {\displaystyle p}numbers) “Linear classifier”.

It choice as the best hyperplane is the one that represents the largest separation, it choose the hyperplane so that the distance from it to the nearest data point on each side is maximized.