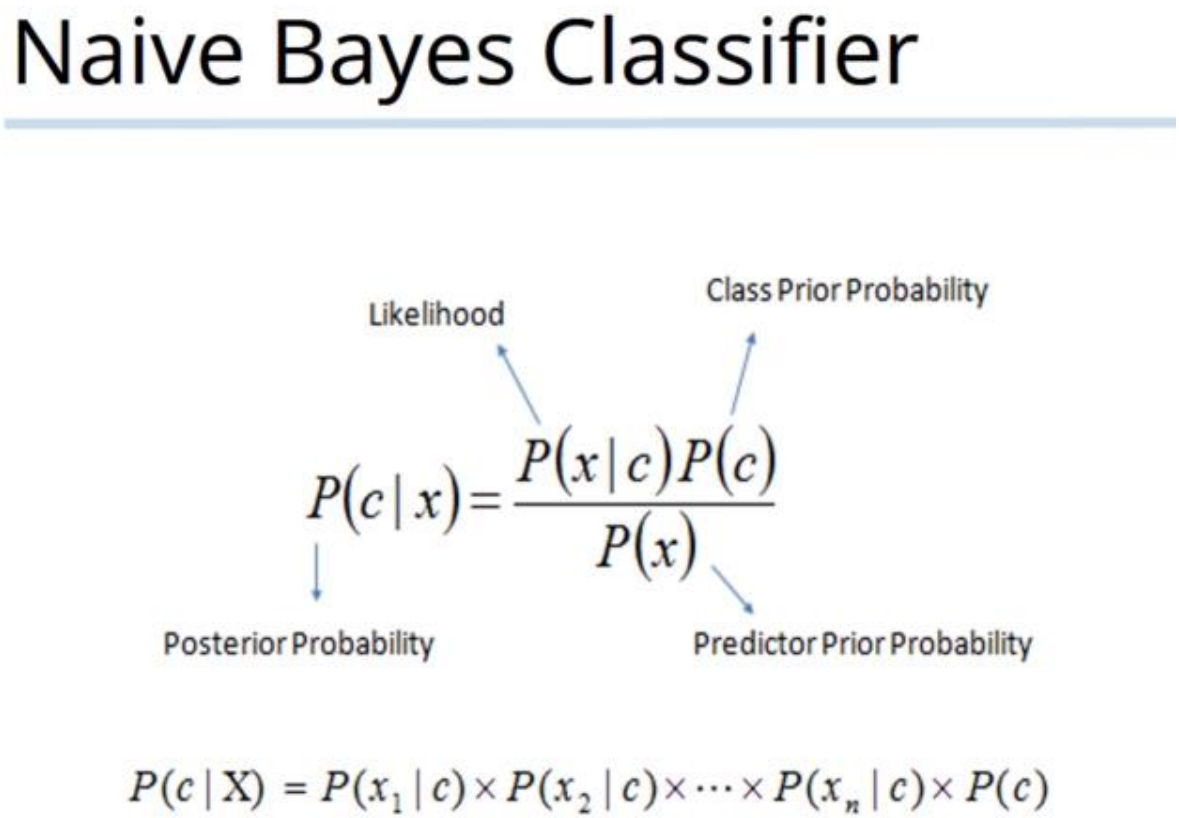
Assignment 2 Part 2

Implementing Naive Bayes Classifier using Spark MapReduce

Naïve-Bayes Algorithm Description:



$$P(c \mid X) = P(x_1 \mid c) \times P(x_2 \mid c) \times \cdots \times P(x_n \mid c) \times P(c)$$

Prior Probabilities:

Prior Positive Probability	0.4988999057062034
Prior Negative Probability	0.5010715204160356

Metrics:

Accuracy	0.85
Precision	0.87
Recall	0.83
F-measure	0.85