

## Assignment 2 Part 2

### Implementing Naive Bayes Classifier using Spark MapReduce

Naïve-Bayes Algorithm Description:

# Naive Bayes Classifier

---

The diagram shows the Naive Bayes formula:  $P(c | x) = \frac{P(x | c)P(c)}{P(x)}$ . Arrows point from labels to the corresponding parts of the formula: 'Likelihood' points to  $P(x | c)$ , 'Class Prior Probability' points to  $P(c)$ , 'Posterior Probability' points to  $P(c | x)$ , and 'Predictor Prior Probability' points to  $P(x)$ .

$$P(c | X) = P(x_1 | c) \times P(x_2 | c) \times \dots \times P(x_n | c) \times P(c)$$

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