

# Learning with Naive Bayes

**Andrew Kirby**

ANDY-KIRBY@LIVE.COM

**Kevin Browder**

BROWDERKEVIN54@GMAIL.COM

**Nathan Stouffer**

NATHANSTOUFFER1999@GMAIL.COM

**Eric Kempf**

ERICKEMPF123@GMAIL.COM

**Editor:**

## Abstract

**Keywords:**

## 1. Introduction

## 2. Problem Statement

The problem that we must solve is a classification problem. Given an input file that contains examples (each example consists of a list of attributes and an associated classification), our task is to implement a learning algorithm that is trained to classify examples. The algorithm we will implement is called Naive Bayes. The performance of our learning algorithm will be evaluated by two metrics of our choosing and 10-fold cross validation. When we have implemented the algorithm, we then perform our experiment. We are tasked with testing whether scrambling values in 10% of the features will affect the performance of Naive Bayes. This effectively eliminates the usefulness of 10% of features in a given data set.

### 2.1 Hypothesis

We predict that scrambling 10% of features will marginally affect performance.  
(Dua and Graff, 2017)

## 3. Algorithm

-It be what it do

## 4. Experimental Design

### 4.1 Set Up

### 4.2 Tuning

Add stuff about bin size, what else?

Experimenting w/ using different attributes

### **4.3 Final Parameters**

- Bin size
- Number of attributes used
- Which attributes used?

## **5. Results**

## **6. Summary**

## **References**

Dheeru Dua and Casey Graff. UCI machine learning repository, 2017. URL <http://archive.ics.uci.edu/ml>.