# Introduction

Many systems nowadays are learning from experience. Think of the dictionary of your smartphone or the home page on your Facebook. These systems are so called interactive learners. In this report we will make our own simple version of such a system. In the following sections the choices we made and the things we are going to make a listed.

# Early Choices

For starters the programming language is the fundament on which you build. For this project Python will be the language at hand. Furthermore we decided that we wanted to grade the different topics on multiple levels. For example for the emails, we state whether or not they or spam, but for jokes, we state how good they are. This leads us to the fact that we will use a multinomial distribution.

# Design and structure of the learner

The design of the learner is written in Python. The choice for this language is pure subjective. In the zip file is specified how the code is structured. In the main file an outline is given about how the code will be executed in basis. The data.py file is responsible for the data aggregation of the test and train data. Train.py is responsible for the actual training of the naïve Bayesian classifier. Classify.py is used to classify documents.

# How to interact

For the interaction we will make a GUI. The GUI has a section for the operation and for development. The operation part has buttons with which you can specify whether you want to read a joke or determine the gender of an author of a blog. There will also be a feedback button. For a joke it can be submitted whether a joke is good or bad or neither, for a blog it can be submitted whether it was a male or a female who wrote it.

The development section has input menu with which you can enter or delete a training set and a button so that training can be initiated.

# Planning

|  |  |  |
| --- | --- | --- |
| Date | What | Who |
| 2015-11-26 | Part A: Code Outline | Joshua |
| 2015-11-26 | Part A: Rapport | Tjeerd |
| 2015-12-07 | Part B: Readme | Joshua |
| 2015-12-07 | Part B: Implement NBC | Tjeerd & Joshua |
| 2015-12-07 | Part B: Build Dataset | Tjeerd |
| 2015-12-07 | Part B: Rapport | Joshua |
| 2015-12-14 | Part C: Build Learner | Tjeerd & Joshua |
| 2015-12-14 | Part C: Build GUI | Tjeerd |
| 2015-12-14 | Part C: Build demonstration | Joshua |
| 2015-12-17 | Part Bonus: Implement χ^2 test | Tjeerd |
| 2015-12-17 | Part Bonus: Generate TestSet | Joshua |
| 2015-12-17 | Part Bonus: Rapport | Tjeerd |
| 2015-12-17 | Part Bonus: Test Classifier of B | Tjeerd |
| 2015-12-17 | Part Bonus: Rapport | Joshua |

# Authors

Tjeerd Jan Heeringa, s1497324

Joshua van Kleef, s1385801