



Machine Learning Exercise (SS 22)

Assignment 2: Naive Bayes

Mojtaba Nayyeri

Mojtaba.Nayyeri@ipvs.uni-stuttgart.de

Alexandra Baier

Alexandra.Baier@ipvs.uni-stuttgart.de

Akram Hosseini

Akram.Hosseini@ipvs.uni-stuttgart.de

Tim Schneider

timphillip.schneider@ipvs.uni-stuttgart.de

Submit your solution in ILIAS as a single PDF file.¹ Make sure to list your full name and immatriculation number at the start of the file. Optionally, you can *additionally* upload source files (e.g. PPTX files). Remember to fill out the exercise slot and exercise presentation polls linked in ILIAS. If you have any questions, feel free to ask them in the exercise forum in ILIAS.

Submission is open until Monday, 09 May 2022, 11:59 AM.

¹Your drawing software probably allows to export as PDF. An alternative option is to use a PDF printer. If you create multiple PDF files, use a merging tool (like [pdfarranger](#)) to combine the PDFs into a single file.



M&M Design and Bayesian Observers

The blue M&M was introduced in 1995. Before then, the color mix in a bag of plain M&Ms was: 20% *Brown*, 30% *Yellow*, 10% *Red*, 20% *Green*, 10% *Orange*, 10% *Tan*. Afterward it was: 20% Blue, 24% Green, 14% Orange, 16% Yellow, 13% Red, 13% Brown. A friend of you has two bags of M&Ms, and she tells you that one is from 1994 and one from 1996. She won't tell you which is which, but she bets that you won't find out.

a) **Task (Prior without information):**

Your friend reaches into her closet and takes out two bags of M&M, putting them side by side on the table. Without further information, what's the probability $p(\text{left} = 1994, \text{right} = 1996)$ for the left bag to be from 1994?

b) **Task (Posterior given evidence):**

Now she gives you one M&M from each bag. The left one is red and right one is brown. What is the probability that the red M&M came from the 1994 bag?

c) **Task (Belief update):**

Obviously, you decided not to eat the M&Ms since they are almost 30 years old. Instead, you put them back and you are shown yet another red one randomly drawn from the left bag. How does this observation change your belief?

d) **Task (Wrong Prior):**

Your friend's mother enters the room, points on the left bag and says that she recognizes an old design only used before 1995 on its backside. This changes your belief to $p(\text{left} = 1994) = 0.9$. How many (consecutive) observations of red M&M independently drawn from the left bag (with replacement) would it need to change your mind ($p(\text{left} = 1994) < 0.5$) again?

Hint: Find a closed-form expression, plot it and solve for its root e.g. with `scipy.optimize.newton`.



Fake News Classification with Naive Bayes (Programming)

1. **Task (Programming)** Please download the Jupyter notebook assignment2.ipynb. Follow the instructions in the Jupyter notebook.



Log-Sum-Exp Trick (Programming)

1. **Task (Programming)** Please download the Jupyter notebook `assignment2.ipynb`. Follow the instructions in the Jupyter notebook.



kNN in High-Dimensional Feature Spaces

1. **Task (Research)** Research and discuss why kNN might fail for high dimensional feature spaces.
2. **Task (Research)** Identify and explain one approach for solving or circumventing this problem.