**Your Name:**

**Your Andrew ID:**

**Homework 2**

# Collaboration and Originality

1. Did you receive help of any kind from anyone in obtaining your data for this assignment (Yes or No)? It is not necessary to describe discussions with the instructor or TA.

If you answered Yes, provide the name(s) of anyone who provided help, and describe the type of help that you received.

1. Did you give help of any kind to anyone in obtaining their data for this assignment (Yes or No)?

If you answered Yes, provide the name(s) of anyone that you helped, and describe the type of help that you provided.

1. Are you the author of every word of your report (Yes or No)?

If you answered No:

* 1. identify the text that you did not write,
  2. explain where it came from, and
  3. explain why you used it.

**Your Name**

**Your Andrew ID**

**Homework 2**

*You must follow this report template. Instructions are given in red italic font throughout the template. Delete the instructions from your report before submitting it.*

*At a high level, this assignment consists of two parts: i) Do an experiment, and ii) analyze and discuss the experiment. The template includes guidance about what to discuss, but consider them general advice, rather than strict instructions. You may discuss other aspects of the experiment that you find useful or interesting. Grading is based more on the quality of your analysis than a specific checklist.*

# Experiment #1: Baselines

## Parameters

*Document your parameter settings. If you changed the parameters, explain why.*

## Results

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Reuters**  **Baseline #1** | | | | | | |
|  | **NB** | | | **SVM** | | |
| **Category** | **P** | **R** | **F1** | **P** | **R** | **F1** |
|  | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
| **Macro Average** |  |  |  |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Reuters**  **Baseline #2** | | | | | | |
|  | **NB** | | | **SVM** | | |
| **Category** | **P** | **R** | **F1** | **P** | **R** | **F1** |
|  | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
| **Macro Average** |  |  |  |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **OHSUMED**  **Baseline #1** | | | | | | |
|  | **NB** | | | **SVM** | | |
| **Category** | **P** | **R** | **F1** | **P** | **R** | **F1** |
|  | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
| **Macro Average** |  |  |  |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **OHSUMED**  **Baseline #2** | | | | | | |
|  | **NB** | | | **SVM** | | |
| **Category** | **P** | **R** | **F1** | **P** | **R** | **F1** |
|  | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
| **Macro Average** |  |  |  |  |  |  |

## Analysis

*Discuss the differences among the two baselines and the two different learning algorithms. Pay particular attention to differences in accuracy and efficiency, and what may have caused them.* *Are categories with more training data easier to learn? If so, is it an important difference?*

# Experiment #2: Feature Selection

## Parameters

*Document your parameter settings. If you changed the parameters, explain why.*

## Results

*Report macro-averaged results for each algorithm (i.e., do not report results for individual categories).*

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Reuters**  **Number of Features** | | | | | | | | | | | | | | |
| **n1=00** | | | **n2=00** | | | **n3=00** | | | **n4=00** | | | **n5=00** | | |
| **P** | **R** | **F1** | **P** | **R** | **F1** | **P** | **R** | **F1** | **P** | **R** | **F1** | **P** | **R** | **F1** |
| **NB** | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| **SVM** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **OHSUMED**  **Number of Features** | | | | | | | | | | | | | | |
| **n1=00** | | | **n2=00** | | | **n3=00** | | | **n4=00** | | | **n5=00** | | |
| **P** | **R** | **F1** | **P** | **R** | **F1** | **P** | **R** | **F1** | **P** | **R** | **F1** | **P** | **R** | **F1** |
| **NB** | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| **SVM** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

## Analysis

*Discuss your results. Are small representations more effective, or are large representations more effective? Does each dataset and/or learning algorithm behave the same way? Does having more features make a difference? If so, is it an important difference? Are the effects similar for small and large classes?*

# Experiment #3: Your Representations

*Report your results in a tabular format similar to what was used for Experiments 1-2. Describe your representations. Discuss your reasons for developing each representation. Discuss how well they worked, and how you interpret the results.*

# Experiment #4: Sentiment Baselines

## Parameters

*Document your parameter settings. If you changed the parameters, explain why.*

## Results

*Report results for each category (Good, Ok, Poor), as well as macro-averaged values.*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Apps for Android** | | | | | | | |
|  | **Category** | **NB** | | | **SVM** | | |
| **P** | **R** | **F1** | **P** | **R** | **F1** |
| **Baseline 1** |  | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
| **Macro Average** |  |  |  |  |  |  |
| **Baseline 2** |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
| **Macro Average** |  |  |  |  |  |  |
| **Custom 1** |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
| **Macro Average** |  |  |  |  |  |  |
| **Custom 2** |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
| **Macro Average** |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Cell Phones and Accessories** | | | | | | | |
|  | **Category** | **NB** | | | **SVM** | | |
| **P** | **R** | **F1** | **P** | **R** | **F1** |
| **Baseline 1** |  | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
| **Macro Average** |  |  |  |  |  |  |
| **Baseline 2** |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
| **Macro Average** |  |  |  |  |  |  |
| **Custom 1** |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
| **Macro Average** |  |  |  |  |  |  |
| **Custom 2** |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
| **Macro Average** |  |  |  |  |  |  |

## Custom Representations

*Describe the custom representations that you used. Explain why you made these choices.*

## Analysis

*Discuss the differences among the three baselines and the two different algorithms. Pay particular attention to differences in accuracy, and what may have caused them.*

# Effectiveness of the Different Representational Choices

*Do you see any trends? Do some techniques consistently work better than others? Do any of the choices affect the stability of results? Do any of the differences seem significant?*

*There is some overlap between this section and observations that you may make in Experiments 1-4. However, in this section you have an opportunity comment on trends that span different experiments, which you can't really do in the earlier sections.*

# Effectiveness of the Different Learning Algorithms

*Do you see any trends? Does one algorithm consistently work better than another? Is there a difference in the stability of the algorithms? Do any of the differences seem significant?*

*There is some overlap between this section and observations that you may make in Experiments 1-4. However, in this section you have an opportunity comment on trends that span different experiments, which you can't really do in the earlier sections.*