CS 440/ECE 448 Artificial Intelligence
Assignment 3: Naive Bayes Classification

Haoen CUI[*], *Guohao (Holden) DOU*[†], *Chuchao LUO*[‡]

November 27, 2017

Contents

[*]Haoen CUI's Email hcui10@illinois.edu
[†]Guohao (Holden) DOU's Email gdou2@illinois.edu
[‡]Chuchao LUO's Email chuchao2@illinois.edu

1 Part 1: Digit Classification

1.1 Single Pixels as Features (For Everybody)

1.1.1 Implementation

We treated this problem as a special case of *pixel group as features* where the pixel groups are simply disjoint and of size 1 by 1. Please see the next section for details.

1.1.2 Smoothing Constant

1.1.3 Classification Rate and Confusion Matrix

1.1.4 Posterior Probabilities: Highest and Lowest

1.1.5 Visualization of Likelihoods and Odds Ratios

1.2 Pixel Groups as Features (For Four-Credit Students)

1.2.1 Accuracy on Test Set

```
sizes_to_run_disj <- c(11, 22, 24, 42, 22)
sizes_to_run_overlap <- c(22, 24, 42, 44, 23, 32, 33)
for (kernel.size in sizes_to_run_disj) {
  i <- kernel.size %/% 10
  j <- kernel.size %% 10
  path.to.file <- paste("img/conf_mat_disj", kernel.size, ".png", sep = "")
  print(path.to.file)
  knitr::include_graphics(path = path.to.file)
}
```

```
## [1] "img/conf_mat_disj11.png"
## [1] "img/conf_mat_disj22.png"
## [1] "img/conf_mat_disj24.png"
## [1] "img/conf_mat_disj42.png"
## [1] "img/conf_mat_disj22.png"
```

```
knitr::include_graphics(path = "img/conf_mat_disj11.png")
```

1.2.2 Trends for Different Feature Sets

1.2.3 Running Time for Different Feature Sets

1.2.3.1 Training

1.2.3.2 Testing

1.3 Extra Credit

1.3.1 Ternary Features

1.3.2 Naive Bayes Classifier on Face Data

---

2 Part 2: Audio Classification

2.1 Binary Classification: Hebrew words of "yes" and "no" (For Everybody)

2.1.1 Implementation

2.1.2 Classification Rate and Confusion Matrix

2.2 Multi-Class Classification: Audio Digits 1-5 Spoken by Four Different Speakers (For Four-Credit Students)

2.2.1 Implementation

2.2.2 Overall Accuracy

2.2.3 Classification Rate and Confusion Matrix

2.3 Extra Credit

2.3.1 Binary Classification on Unsegmented Data

2.3.2 Alternative Method (RNN) on XXX Data

2.3.3 Average-Column Method on Hebrew Yes-No Corpus

---

3 Statement of Individual Contribution

Table 1: Statement of Individual Contribution

	NetID	Contribution
Haoen CUI	hcui10	visualization, report, and ideas generation
Guohao DOU	gdou2	part 1 (algorithm design and programming) and ideas generation
Chuchao LUO	chuchao2	part 2 (algorithm design and programming) and ideas generation