CS 66: Machine Learning

Prof. Sara Mathieson

Spring 2019
• **Lab 3 due Thursday**
  – If you are finished, you do not need to come to lab today (but you are welcome to start on Lab 4)
  – If you are “finished” but don’t understand something, please come to lab and ask!

• **Midterm 1 Feb 27** (in lab)
  – Pick up a study guide!
  – You may use a 1-page (front and back) “cheat-sheet”, but no other resources

• **Lab 4 due March 8** (Friday before spring break)
  – Will run partner script right after class
  – Let me know ASAP of partner preferences!
Outline for February 20

• Naïve Bayes

• Bayes Theorem

• Confusion Matrices
Bayes Rule

\[ P(A, B) = P(A)P(B \mid A) = P(B)P(A \mid B) \]

\[ \Rightarrow P(A)P(B \mid A) = P(B)P(A \mid B) \]

\[ \Rightarrow P(B \mid A) = \frac{P(B)P(A \mid B)}{P(A)} \]

Naive Bayes Assumption

\[ \Rightarrow \text{feature } j \text{ is independent of all other features} \]

\[ \text{give class label} \]
\[ P(x_1, x_2, \ldots, x_p | y=k) \]
\[ \approx \prod_{j=1}^p P(x_j | y=k) \]

Model:
\[ P(y=k | x) \propto P(y=k) \prod_{j=1}^p P(x_j | y=k) \]

Prediction:
\[ \hat{y} = \text{argmax}_k P(y=k) \prod_{j=1}^p P(x_j | y=k) \]

\[ \theta_k, \theta_{j,v,r} \text{ plug in} \]
<table>
<thead>
<tr>
<th>X</th>
<th>f₁</th>
<th>f₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>X₁</td>
<td>pos</td>
<td>neg</td>
</tr>
<tr>
<td>X₂</td>
<td>pos</td>
<td>pos</td>
</tr>
<tr>
<td>X₃</td>
<td>pos</td>
<td>neg</td>
</tr>
<tr>
<td>X₄</td>
<td>neg</td>
<td>neg</td>
</tr>
<tr>
<td>X₅</td>
<td>pos</td>
<td>neg</td>
</tr>
<tr>
<td>X₆</td>
<td>neg</td>
<td>neg</td>
</tr>
<tr>
<td>X₇</td>
<td>neg</td>
<td>pos</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Y</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>X₁</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>X₂</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>X₃</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>X₄</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

**Estimate** $p(y=k)$

$$θ_k = \frac{N_k + 1}{n + K}$$

$LäPlace$ counts

$N_k = \#$ of datapoints with class $k$

$N₁ = 3$, $N₂ = 4$

Handout 4 example
$N_1 = 3, \quad N_2 = 4$

$\Theta_1 = \frac{3 + 1}{7 + 2} = \frac{4}{9}$

$\Theta_2 = \frac{4 + 1}{7 + 2} = \frac{5}{9}$

add to 1

Handout 4, Question 1
\[ \Theta_1 + \Theta_2 + \ldots + \Theta_K = \frac{(N_1 + 1) + (N_2 + 1) \cdots (N_K + 1)}{n + \varepsilon} = \sum_{k=1}^{K} \frac{N_k + K}{n + K} \]

\[ \Theta_{j,v,k} = \frac{N_{j,v,k} + 1}{N_k + |f_j|} \]

# of possible values feature j can take on

estimate for 

\[ P(x_j = v | y = k) \]
## Confusion Matrix

### Handout 4, Q2

<table>
<thead>
<tr>
<th></th>
<th>$Y = 1$</th>
<th>$Y = 2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$pos$</td>
<td>( \frac{1+1}{3+2} )</td>
<td>( \frac{3+1}{4+2} )</td>
</tr>
<tr>
<td>$neg$</td>
<td>( \frac{2+1}{3+2} )</td>
<td>( \frac{1+1}{4+2} )</td>
</tr>
</tbody>
</table>

Add to:

- 1

### Confusion Matrix

![Confusion Matrix Diagram]

*Pred label*

![Confusion Matrix Diagram]

*True label*

1 1 2

1 1 1

diagional to be high!