Classification: given inputs $x$, predict labels (classes) $y$

Naïve Bayes

$$P(Y|F_0,0 \ldots F_{15,15}) \propto P(Y) \prod_{i,j} P(F_{i,j}|Y)$$
Parameter Estimation
Parameter Estimation with Maximum Likelihood

- **Estimating the distribution of a random variable**
  - E.g.: for each outcome x, look at the *empirical rate* of that value:
    \[ P_{\text{ML}}(x) = \frac{\text{count}(x)}{\text{total samples}} \]
    \[ P_{\text{ML}}(r) = 2/3 \]
  - This is the estimate of the parameters that maximizes the *likelihood of the data*

\[ L(x, \theta) = \prod_{i} P_{\theta}(x_i) = \theta \cdot \theta \cdot (1 - \theta) \]

\[ P_{\theta}(x = \text{red}) = \theta \]
\[ P_{\theta}(x = \text{blue}) = 1 - \theta \]
Parameter Estimation with Maximum Likelihood

- **Data:** Observed set $D$ of $\alpha_H$ Heads and $\alpha_T$ Tails
- **Hypothesis space:** Binomial distributions
- **Learning:** finding $\theta$ is an optimization problem
  - What’s the objective function?
    \[ P(D \mid \theta) = \theta^{\alpha_H} (1 - \theta)^{\alpha_T} \]
- **MLE:** Choose $\theta$ to maximize probability of $D$
  \[ \hat{\theta} = \arg \max_\theta P(D \mid \theta) \]
  \[ = \arg \max_\theta \ln P(D \mid \theta) \]
Parameter Estimation with Maximum Likelihood

\[ \hat{\theta} = \arg \max_{\theta} \ln P(\mathcal{D} \mid \theta) \]

\[ = \arg \max_{\theta} \ln \theta^{\alpha_H} (1 - \theta)^{\alpha_T} \]

- Set derivative to zero, and solve!

\[ \frac{d}{d\theta} \ln P(\mathcal{D} \mid \theta) = \frac{d}{d\theta} \left[ \ln \theta^{\alpha_H} (1 - \theta)^{\alpha_T} \right] \]

\[ = \frac{d}{d\theta} \left[ \alpha_H \ln \theta + \alpha_T \ln(1 - \theta) \right] \]

\[ = \alpha_H \frac{d}{d\theta} \ln \theta + \alpha_T \frac{d}{d\theta} \ln(1 - \theta) \]

\[ = \frac{\alpha_H}{\theta} - \frac{\alpha_T}{1 - \theta} = 0 \]

\[ \hat{\theta}_{MLE} = \frac{\alpha_H}{\alpha_H + \alpha_T} \]
Parameter Estimation with Maximum Likelihood

- How do we estimate the conditional probability tables?
  - Maximum Likelihood, which corresponds to counting

- Need to be careful though … let’s see what can go wrong..
Underfitting and Overfitting
Example: Overfitting

\[ P(\text{features}, C = 2) \]

\[ P(C = 2) = 0.1 \]

\[ P(\text{on}|C = 2) = 0.8 \]

\[ P(\text{on}|C = 2) = 0.1 \]

\[ P(\text{off}|C = 2) = 0.1 \]

\[ P(\text{on}|C = 2) = 0.01 \]

\[ P(\text{features}, C = 3) \]

\[ P(C = 3) = 0.1 \]

\[ P(\text{on}|C = 3) = 0.8 \]

\[ P(\text{on}|C = 3) = 0.9 \]

\[ P(\text{off}|C = 3) = 0.7 \]

\[ P(\text{on}|C = 3) = 0.0 \]

\textit{2 wins!!}
Overfitting

Degree 15 polynomial
Training and Testing
Important Concepts

- Data: labeled instances, e.g. emails marked spam/ham
  - Training set
  - Held out set
  - Test set
- Features: attribute-value pairs which characterize each x
- Experimentation cycle
  - Learn parameters (e.g. model probabilities) on training set
  - (Tune hyperparameters on held-out set)
  - Compute accuracy on test set
  - Very important: never “peek” at the test set!
- Evaluation
  - Accuracy: fraction of instances predicted correctly
- Overfitting and generalization
  - Want a classifier which does well on test data
  - Overfitting: fitting the training data very closely, but not generalizing well
  - Underfitting: fits the training set poorly
Generalization and Overfitting

- Relative frequency parameters will **overfit** the training data!
  - Just because we never saw a 3 with pixel (15,15) on during training doesn’t mean we won’t see it at test time
  - Unlikely that every occurrence of “minute” is 100% spam
  - Unlikely that every occurrence of “seriously” is 100% ham
  - What about all the words that don’t occur in the training set at all?
  - In general, we can’t go around giving unseen events zero probability

- As an extreme case, imagine using the entire email as the only feature
  - Would get the training data perfect (if deterministic labeling)
  - Wouldn’t **generalize** at all
  - Just making the bag-of-words assumption gives us some generalization, but isn’t enough

- To generalize better: we need to **smooth** or **regularize** the estimates
Smoothing
Unseen Events
Laplace Smoothing

- **Laplace’s estimate:**
  - Pretend you saw every outcome once more than you actually did

  \[ P_{LAP}(x) = \frac{c(x) + 1}{\sum_x [c(x) + 1]} \]

  \[ = \frac{c(x) + 1}{N + |X|} \]

- Can derive this estimate with *Dirichlet priors* (see cs281a)
Laplace Smoothing

- **Laplace’s estimate (extended):**
  - Pretend you saw every outcome k extra times
  
  \[ P_{LAP,k}(x) = \frac{c(x) + k}{N + k|X|} \]
  
  - What’s Laplace with k = 0?
  - k is the **strength** of the prior

- **Laplace for conditionals:**
  - Smooth each condition independently:
  
  \[ P_{LAP,k}(x|y) = \frac{c(x, y) + k}{c(y) + k|X|} \]
Formal Derivation

- Relative frequencies are the maximum likelihood estimates

\[
\theta_{ML} = \arg \max_\theta P(X|\theta) \\
= \arg \max_\theta \prod_i P_\theta(X_i)
\]

\[
P_{ML}(x) = \frac{\text{count}(x)}{\text{total samples}}
\]

- Another option is to consider the most likely parameter value given the data

\[
\theta_{MAP} = \arg \max_\theta P(\theta|X) \\
= \arg \max_\theta P(X|\theta)P(\theta)/P(X) \\
= \arg \max_\theta P(X|\theta)P(\theta)
\]

“right” choice of \(P(\theta)\)  
-> Laplace estimates
Now we’ve got two kinds of unknowns
- Parameters: the probabilities $P(X|Y), P(Y)$
- Hyperparameters: e.g. the amount / type of smoothing to do, $k, \alpha$

What should we learn where?
- Learn parameters from training data
- Tune hyperparameters on different data
  - Why?
- For each value of the hyperparameters, train and test on the held-out data
- Choose the best value and do a final test on the test data
Practical Tip: Baselines

- **First step: get a baseline**
  - Baselines are very simple “straw man” procedures
  - Help determine how hard the task is
  - Help know what a “good” accuracy is

- **Weak baseline: most frequent label classifier**
  - Gives all test instances whatever label was most common in the training set
  - E.g. for spam filtering, might label everything as ham
  - Accuracy might be very high if the problem is skewed
  - E.g. calling everything “ham” gets 66%, so a classifier that gets 70% isn’t very good…

- For real research, usually use previous work as a (strong) baseline
Summary

- Bayes rule lets us do diagnostic queries with causal probabilities
- The naïve Bayes assumption takes all features to be independent given the label
- We can build classifiers out of a naïve Bayes model using training data
- Smoothing estimates is important in real systems
Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

$x$  $f(x)$  $y$

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Do you want free printr cartriges? Why pay more when you can get them ABSOLUTELY FREE! Just

Hello,
Some (Simplified) Biology

- Very loose inspiration: human neurons
Linear Classifiers

- Inputs are feature values
- Each feature has a weight
- Sum is the activation

\[
\text{activation}_w(x) = \sum_i^{} w_i \cdot f_i(x) = w \cdot f(x)
\]

- If the activation is:
  - Positive, output +1
  - Negative, output -1
Weights

- Binary case: compare features to a weight vector
- Learning: figure out the weight vector from examples

\[ w \cdot f(x) \text{ positive means the positive class} \]
Decision Rules
Binary Decision Rule

- In the space of feature vectors
  - Examples are points
  - Any weight vector is a hyperplane
  - One side corresponds to $Y=+1$
  - Other corresponds to $Y=-1$

$$w$$

<table>
<thead>
<tr>
<th>BIAS</th>
<th>-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>free</td>
<td>4</td>
</tr>
<tr>
<td>money</td>
<td>2</td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>
In the space of feature vectors
- Examples are points
- Any weight vector is a hyperplane
- One side corresponds to $Y=+1$
- Other corresponds to $Y=-1$

$w$

BIAS : -3
free : 4
money : 2
...

$0 \leq f \leq 1$

$-1 = \text{HAM}$

$+1 = \text{SPAM}$

$f \cdot w = 0$
Weight Updates
Learning: Binary Perceptron

- Start with weights = 0
- For each training instance:
  - Classify with current weights
  - If correct (i.e., y=y*), no change!
  - If wrong: adjust the weight vector
Learning: Binary Perceptron

- Start with weights $= 0$
- For each training instance:
  - Classify with current weights
    \[ y = \begin{cases} 
    +1 & \text{if } w \cdot f(x) \geq 0 \\
    -1 & \text{if } w \cdot f(x) < 0 
    \end{cases} \]
  - If correct (i.e., $y=y^*$), no change!
  - If wrong: adjust the weight vector by adding or subtracting the feature vector. Subtract if $y^*$ is -1.

\[ w = w + y^* \cdot f \]

Before: $wf$
After: $wf + y^*ff$
\[ ff \geq 0 \]
Examples: Perceptron

- Separable Case
If we have multiple classes:
- A weight vector for each class: $w_y$
- Score (activation) of a class $y$: $w_y \cdot f(x)$
- Prediction highest score wins

$$y = \arg \max_y w_y \cdot f(x)$$

Binary = multiclass where the negative class has weight zero
Learning: Multiclass Perceptron

- Start with all weights = 0
- Pick up training examples one by one
- Predict with current weights
  \[ y = \arg \max_y w_y \cdot f(x) \]
- If correct, no change!
- If wrong: lower score of wrong answer, raise score of right answer
  \[ w_y = w_y - f(x) \]
  \[ w_{y^*} = w_{y^*} + f(x) \]
Properties of Perceptrons

- **Separability**: true if some parameters get the training set perfectly correct

- **Convergence**: if the training is separable, perceptron will eventually converge (binary case)
Problems with the Perceptron

- Noise: if the data isn’t separable, weights might thrash
  - Averaging weight vectors over time can help (averaged perceptron)

- Mediocre generalization: finds a “barely” separating solution

- Overtraining: test / held-out accuracy usually rises, then falls
  - Overtraining is a kind of overfitting
Example: Multiclass Perceptron

“win the vote” \[ [1 \ 1 \ 0 \ 1 \ 1] \]

“win the election” \[ [1 \ 1 \ 0 \ 0 \ 1] \]

“win the game” \[ [1 \ 1 \ 1 \ 0 \ 1] \]

<table>
<thead>
<tr>
<th>(_w^{SPORTS})</th>
<th>1</th>
<th>-2</th>
<th>-2</th>
<th>(_w^{POLITICS})</th>
<th>0</th>
<th>3</th>
<th>3</th>
<th>(_w^{TECH})</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIAS</td>
<td>1</td>
<td></td>
<td></td>
<td>BIAS</td>
<td>0</td>
<td>1</td>
<td></td>
<td>BIAS</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>win</td>
<td>0</td>
<td>-1</td>
<td></td>
<td>win</td>
<td>0</td>
<td>1</td>
<td></td>
<td>win</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>game</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>game</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>game</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>vote</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>vote</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>vote</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>the</td>
<td>0</td>
<td>-1</td>
<td></td>
<td>the</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>the</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>