CPSC 425: Computer Vision

Lecture 17: Classification
Classification

Problem:
Assign new observations into one of a fixed set of categories (classes)

Key Idea(s):
Build a model of data in a given category based on observations of instances in that category
Classification

(assume given set of discrete labels) \{dog, cat, truck, plane, \ldots\}

\[ \text{cat} \]
Classification

What the computer sees

image classification

82% cat
15% dog
2% hat
1% mug
Classification

A **classifier** is a procedure that accepts as input a set of features and outputs a class **label**.

Classifiers can be binary (face vs. not-face) or multi-class (cat, dog, horse, ...).

We build a classifier using a **training set** of labelled examples \( \{(x_i, y_i)\} \), where each \( x_i \) is a feature vector and each \( y_i \) is a class label.

Given a previously unseen observation, we use the classifier to predict its class label.
Classification

— Collect a database of images with labels
— Use ML to train an image classifier
— Evaluate the classifier on test images
Example 1: A Classification Problem

Categorize images of fish — “Atlantic salmon” vs “Pacific salmon”

Use **features** such as length, width, lightness, fin shape & number, mouth position, etc.

Given a previously unobserved image of a salmon, use the learned classifier to guess whether it is an Atlantic or Pacific salmon.

**Figure credit:** Duda & Hart
**Example 2: Real Classification Problem**

**SUN Dataset**
- 131K images
- 908 *scene* categories

<table>
<thead>
<tr>
<th>Indoor</th>
<th>Shopping and dining</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>auto showroom</td>
</tr>
<tr>
<td>Outdoor natural</td>
<td>workplace (office building, factory, lab, etc.)</td>
</tr>
<tr>
<td></td>
<td>bakery kitchen</td>
</tr>
<tr>
<td>Outdoor man-made</td>
<td>home or hotel</td>
</tr>
<tr>
<td></td>
<td>bakery shop</td>
</tr>
<tr>
<td>Transportation (vehicle interiors, stations, etc.)</td>
<td>bank indoor</td>
</tr>
<tr>
<td></td>
<td>bank vault</td>
</tr>
<tr>
<td>Sports and leisure</td>
<td>banquet hall</td>
</tr>
<tr>
<td>Cultural (art, education, religion, military, law, politics, etc.)</td>
<td>bar</td>
</tr>
</tbody>
</table>
Example 3: Real Classification Problem

ImageNet Dataset

- 14 Million images

- 21K object categories
Bayes Rule (Review and Definitions)

Let $c$ be the class label and let $x$ be the measurement (i.e., evidence)

$$P(c|x) = \frac{P(x|c)p(c)}{P(x)}$$

- $P(c|x)$: posterior probability
- $P(x|c)$: class-conditional probability (a.k.a. likelihood)
- $p(c)$: prior probability
- $P(x)$: unconditional probability (a.k.a. marginal likelihood)
Bayes Rule (Review and Definitions)

Let \( c \) be the **class label** and let \( x \) be the **measurement** (i.e., evidence)

**Simple** case:
- binary classification; i.e., \( c \in \{1, 2\} \)
- features are 1D; i.e., \( x \in \mathbb{R} \)

\[
P(c|x) = \frac{P(x|c)p(c)}{P(x)}
\]

**General** case:
- multi-class; i.e., \( c \in \{1, ..., 1000\} \)
- features are high-dimensional; i.e., \( x \in \mathbb{R}^{2,000+} \)
Example: Discrete Bayes Classifier

Assume we have two classes: \( c_1 = \text{male} \quad c_2 = \text{female} \)

We have a person who’s gender we don’t know, who’s name is Drew.
Example: Discrete Bayes Classifier

Assume we have two classes: \( c_1 = \text{male} \quad c_2 = \text{female} \)

We have a person who’s gender we don’t know, who’s name is \textit{drew}
Example: Discrete Bayes Classifier

Assume we have two classes: \( c_1 = \text{male} \quad c_2 = \text{female} \)

We have a person who’s gender we don’t know, who’s name is *drew*

Classifying *drew* as being male or female is equivalent to asking is it more probable that *drew* is male or female, i.e. which is greater

\[
p(\text{male}|\text{drew}) \quad p(\text{female}|\text{drew})
\]

Example from: Eamonn Keogh
Example: Discrete Bayes Classifier

Assume we have two classes: \( c_1 = \text{male} \) \( c_2 = \text{female} \)

We have a person who’s gender we don’t know, who’s name is \( \text{drew} \)

Classifying \( \text{drew} \) as being male or female is equivalent to asking is it more probable that \( \text{drew} \) is male or female, i.e. which is greater \( p(\text{male}|\text{drew}) \) \( p(\text{female}|\text{drew}) \)

\[
\begin{align*}
p(\text{male}|\text{drew}) &= \frac{p(\text{drew}|\text{male})p(\text{male})}{p(\text{drew})} \\
p(\text{female}|\text{drew}) &= \frac{p(\text{drew}|\text{female})p(\text{female})}{p(\text{drew})}
\end{align*}
\]

Example from: Eamonn Keogh
Example: Discrete Bayes Classifier

\[ p(\text{male}|drew) = \frac{p(drew|\text{male})p(\text{male})}{p(drew)} \]
Example: Discrete Bayes Classifier

\[ p(\text{male}) = \]

\[ p(\text{drew}|\text{male}) = \]

\[ p(\text{drew}) = \]

\[ p(\text{male}|\text{drew}) = \frac{p(\text{drew}|\text{male})p(\text{male})}{p(\text{drew})} \]

<table>
<thead>
<tr>
<th>Name</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drew</td>
<td>Male</td>
</tr>
<tr>
<td>Claudia</td>
<td>Female</td>
</tr>
<tr>
<td>Drew</td>
<td>Female</td>
</tr>
<tr>
<td>Drew</td>
<td>Female</td>
</tr>
<tr>
<td>Alberto</td>
<td>Male</td>
</tr>
<tr>
<td>Karin</td>
<td>Female</td>
</tr>
<tr>
<td>Nina</td>
<td>Female</td>
</tr>
<tr>
<td>Sergio</td>
<td>Male</td>
</tr>
</tbody>
</table>
Example: Discrete Bayes Classifier

\[
p(\text{male}) = \frac{3}{8}
\]

\[
p(\text{drew}|\text{male}) = 
\]

\[
p(\text{drew}) = 
\]

\[
p(\text{male}|\text{drew}) = \frac{p(\text{drew}|\text{male})p(\text{male})}{p(\text{drew})}
\]

Name | Gender
--- | ---
Drew | Male
Claudia | Female
Drew | Female
Drew | Female
Alberto | Male
Karin | Female
Nina | Female
Sergio | Male

Example from: Eamonn Keogh
Example: Discrete Bayes Classifier

\[ p(\text{male}) = \frac{3}{8} \]

\[ p(\text{drew}|\text{male}) = \frac{1}{3} \]

\[ p(\text{drew}) = \]

\[ p(\text{male}|\text{drew}) = \frac{p(\text{drew}|\text{male})p(\text{male})}{p(\text{drew})} \]

<table>
<thead>
<tr>
<th>Name</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drew</td>
<td>Male</td>
</tr>
<tr>
<td>Claudia</td>
<td>Female</td>
</tr>
<tr>
<td>Drew</td>
<td>Female</td>
</tr>
<tr>
<td>Drew</td>
<td>Female</td>
</tr>
<tr>
<td>Alberto</td>
<td>Male</td>
</tr>
<tr>
<td>Karin</td>
<td>Female</td>
</tr>
<tr>
<td>Nina</td>
<td>Female</td>
</tr>
<tr>
<td>Sergio</td>
<td>Male</td>
</tr>
</tbody>
</table>

Example from: Eamonn Keogh
Example: Discrete Bayes Classifier

\[ p(\text{male}) = \frac{3}{8} \]

\[ p(\text{drew}|\text{male}) = \frac{1}{3} \]

\[ p(\text{drew}) = \frac{3}{8} \]

\[ p(\text{male}|\text{drew}) = \frac{p(\text{drew}|\text{male})p(\text{male})}{p(\text{drew})} \]

<table>
<thead>
<tr>
<th>Name</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drew</td>
<td>Male</td>
</tr>
<tr>
<td>Claudia</td>
<td>Female</td>
</tr>
<tr>
<td>Drew</td>
<td>Female</td>
</tr>
<tr>
<td>Albert</td>
<td>Male</td>
</tr>
<tr>
<td>Karin</td>
<td>Female</td>
</tr>
<tr>
<td>Nina</td>
<td>Female</td>
</tr>
<tr>
<td>Sergio</td>
<td>Male</td>
</tr>
</tbody>
</table>

Example from: Eamonn Keogh
**Example:** Discrete Bayes Classifier

\[
p(\text{male}) = \frac{3}{8}
\]

\[
p(\text{drew}|\text{male}) = \frac{1}{3}
\]

\[
p(\text{drew}) = \frac{3}{8}
\]

\[
p(\text{male}|\text{drew}) = \frac{p(\text{drew}|\text{male})p(\text{male})}{p(\text{drew})} = \frac{3}{8} \times \frac{3}{8} = 0.125
\]

<table>
<thead>
<tr>
<th>Name</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drew</td>
<td>Male</td>
</tr>
<tr>
<td>Claudia</td>
<td>Female</td>
</tr>
<tr>
<td>Drew</td>
<td>Female</td>
</tr>
<tr>
<td>Drew</td>
<td>Female</td>
</tr>
<tr>
<td>Alberto</td>
<td>Male</td>
</tr>
<tr>
<td>Karin</td>
<td>Female</td>
</tr>
<tr>
<td>Nina</td>
<td>Female</td>
</tr>
<tr>
<td>Sergio</td>
<td>Male</td>
</tr>
</tbody>
</table>
Example: Discrete Bayes Classifier

\[
p(\text{male}) = \frac{3}{8} \quad \quad p(\text{female}) = \frac{5}{8}
\]

\[
p(\text{drew}|\text{male}) = \frac{1}{3} \quad \quad p(\text{drew}|\text{female}) = \frac{2}{5}
\]

\[
p(\text{drew}) = \frac{3}{8}
\]

\[
p(\text{male}|\text{drew}) = \frac{p(\text{drew}|\text{male})p(\text{male})}{p(\text{drew})} = 0.125
\]

\[
p(\text{female}|\text{drew}) = \frac{p(\text{drew}|\text{female})p(\text{female})}{p(\text{drew})} = 0.25
\]
Bayes Rule (Review and Definitions)

Let $c$ be the **class label** and let $x$ be the **measurement** (i.e., evidence)

**Simple** case:
- binary classification; i.e., $c \in \{1, 2\}$
- features are 1D; i.e., $x \in \mathbb{R}$

$$P(c|x) = \frac{P(x|c)p(c)}{P(x)}$$

**General** case:
- multi-class; i.e., $c \in \{1, ..., 1000\}$
- features are high-dimensional; i.e., $x \in \mathbb{R}^{2,000+}$
Bayes’ Risk

Some errors may be inevitable: the minimum risk (shaded area) is called the Bayes’ risk

Forsyth & Ponce (2nd ed.) Figure 15.1
**Discriminative vs. Generative**

Finding a **decision boundary** is not the same as modeling a **conditional density** — while a normal density here is a poor fit to $P(1|x)$, the quality of the classifier depends only on how well the boundary is positioned.

Forsyth & Ponce (2nd ed.) Figure 15.5
Discriminative vs. Generative

Finding a decision boundary is not the same as modeling a conditional density — while a normal density here is a poor fit to $P(1|x)$, the quality of the classifier depends only on how well the boundary is positioned.

Forsyth & Ponce (2nd ed.) Figure 15.5