Machine Learning & Algorithmic Bias

CS16: Introduction to Data Structures & Algorithms
Seny Kamara - Spring 2017
Outline

- Motivation
- Supervised learning
- Decision Trees
- Algorithmic Bias
Machine Learning

- Algorithms that
  - learn from data to make predictions
- Allows us to design algorithms
  - that predict the future (e.g., picking stocks)
  - when we don’t know how (e.g., facial recognition)
Classes of ML

‣ Supervised learning
  ‣ learn a function from labeled training data

‣ Unsupervised learning
  ‣ find patterns in data without training data

‣ Reinforcement learning
  ‣ design algorithm that improves with positive and negative feedback
Supervised Learning

- Learn a function from training data
- Given a training set
  - \((x_1, y_1), \ldots, (x_n, y_n) \in X \times Y\)
  - produced by some function \(f\)
  - for all \(i\), \(y_i = f(x_i)\) for some function \(f\)
- Find function \(h\) from a hypothesis space \(H\) that
  - is consistent with \(f\) on training set
  - does well on new inputs
Supervised Learning

- Classification
  - If outputs $y_i$ are from a finite set $\mathbf{Y}$
  - ex: spam/not spam, rainy/sunny/cloudy

- Regression
  - If outputs are from the reals
  - ex: temperature
Making Decisions

- Given a training set
  - \((x_1, y_1), \ldots, (x_n, y_n) \in X \times Y\)
  - such that \(y_i = f(x_i)\)
- for some function \(f\) with form
  - \(f : F_1 \times \ldots \times F_n \rightarrow X\)
  - where \(F_i\) are features
  - \(X\) is a decision
- Find a function \(h\) from a hypothesis space \(H\)
  - that agrees with \(f\) and does well on future decisions
Example: Waiting for a Table?

- **Features**
  - $F_1$: Alternatives = {Yes, No}
  - $F_2$: Bar = {Yes, No}
  - $F_3$: Fri/Sat = {Yes, No}
  - $F_4$: Hungry = {Yes, No}
  - $F_5$: Patrons = {None, Some, Full}
  - $F_6$: Price = {$, $$, $$$}
  - $F_7$: Raining = {Yes, No}
  - $F_8$: Reservation = {Yes, No}
  - $F_9$: Type = {French, Italian, Thai, Burger}
  - $F_{10}$: Wait = {10-30, 30-60, >60}
  - Decision: $X = \{\text{Yes, No}\}$
## Training Data

<table>
<thead>
<tr>
<th>Example</th>
<th>Alt</th>
<th>Bar</th>
<th>Fri</th>
<th>Hun</th>
<th>Pat</th>
<th>Price</th>
<th>Rain</th>
<th>Res</th>
<th>Type</th>
<th>Est</th>
<th>Goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Some</td>
<td>$$$</td>
<td>No</td>
<td>Yes</td>
<td>French</td>
<td>0–10</td>
<td>$y_1 = Yes$</td>
</tr>
<tr>
<td>$x_2$</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Full</td>
<td>$</td>
<td>No</td>
<td>No</td>
<td>Thai</td>
<td>30–60</td>
<td>$y_2 = No$</td>
</tr>
<tr>
<td>$x_3$</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Some</td>
<td>$</td>
<td>No</td>
<td>No</td>
<td>Burger</td>
<td>0–10</td>
<td>$y_3 = Yes$</td>
</tr>
<tr>
<td>$x_4$</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Full</td>
<td>$</td>
<td>Yes</td>
<td>No</td>
<td>Thai</td>
<td>10–30</td>
<td>$y_4 = Yes$</td>
</tr>
<tr>
<td>$x_5$</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Full</td>
<td>$$$</td>
<td>No</td>
<td>Yes</td>
<td>French</td>
<td>&gt;60</td>
<td>$y_5 = No$</td>
</tr>
<tr>
<td>$x_6$</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Some</td>
<td>$$</td>
<td>Yes</td>
<td>Yes</td>
<td>Italian</td>
<td>0–10</td>
<td>$y_6 = Yes$</td>
</tr>
<tr>
<td>$x_7$</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>None</td>
<td>$</td>
<td>Yes</td>
<td>No</td>
<td>Burger</td>
<td>0–10</td>
<td>$y_7 = No$</td>
</tr>
<tr>
<td>$x_8$</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Some</td>
<td>$$</td>
<td>Yes</td>
<td>Yes</td>
<td>Thai</td>
<td>0–10</td>
<td>$y_8 = Yes$</td>
</tr>
<tr>
<td>$x_9$</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Full</td>
<td>$</td>
<td>Yes</td>
<td>No</td>
<td>Burger</td>
<td>&gt;60</td>
<td>$y_9 = No$</td>
</tr>
<tr>
<td>$x_{10}$</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Full</td>
<td>$$$</td>
<td>No</td>
<td>Yes</td>
<td>Italian</td>
<td>10–30</td>
<td>$y_{10} = No$</td>
</tr>
<tr>
<td>$x_{11}$</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>None</td>
<td>$</td>
<td>No</td>
<td>No</td>
<td>Thai</td>
<td>0–10</td>
<td>$y_{11} = No$</td>
</tr>
<tr>
<td>$x_{12}$</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Full</td>
<td>$</td>
<td>No</td>
<td>No</td>
<td>Burger</td>
<td>30–60</td>
<td>$y_{12} = Yes$</td>
</tr>
</tbody>
</table>

S. Russel & P. Norvig. Artificial Intelligence - A Modern Approach
Decision Trees

- Simple and useful representation of functions
  - $f : F_1 \times \ldots \times F_n \rightarrow X$
- In our problem
  - $h$ is a decision tree
  - $H$ is the space of all decision trees
Making Decisions

- Given a training set
  - \((x_1, y_1), \ldots, (x_n, y_n) \in X \times Y\)
  - such that \(y = f(x_i)\)
- for some function \(f\) with form
  - \(f : F_1 \times \ldots \times F_n \rightarrow X\)
  - where \(F_i\) are features
  - \(X\) is a decision
- Find a function \(h\) from a hypothesis space \(H\)
  - that agrees with \(f\) and does well on future decisions
Decision Tree Example

S. Russel & P. Norvig. Artificial Intelligence - A Modern Approach
Learning a Decision Tree

Data

Learn

S. Russel & P. Norvig. Artificial Intelligence - A Modern Approach
What is a Good Decision Tree?

- Three properties
  - consistent with training data
  - performs well on future inputs
  - as small as possible

- How can we find a small decision tree?
  - there are $\Omega(2^{2^n})$ possible decision trees
  - so brute force is not possible
ID3 Algorithm

<table>
<thead>
<tr>
<th>Example</th>
<th>Alt</th>
<th>Bar</th>
<th>Fri</th>
<th>Hun</th>
<th>Pat</th>
<th>Price</th>
<th>Rain</th>
<th>Res</th>
<th>Type</th>
<th>Est</th>
<th>WillWait</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Some</td>
<td>$$$$</td>
<td>No</td>
<td>Yes</td>
<td>French</td>
<td>0-10</td>
<td>y1 = Yes</td>
</tr>
<tr>
<td>X2</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Full</td>
<td>$</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Burger</td>
<td>0-10</td>
</tr>
<tr>
<td>X3</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Some</td>
<td>$</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Thai</td>
<td>10-30</td>
</tr>
<tr>
<td>X4</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Full</td>
<td>$</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Thai</td>
<td>10-30</td>
</tr>
<tr>
<td>X5</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Full</td>
<td>$$$$</td>
<td>No</td>
<td>Yes</td>
<td>French</td>
<td>&gt;60</td>
<td>y5 = No</td>
</tr>
<tr>
<td>X6</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Some</td>
<td>$</td>
<td>Yes</td>
<td>Yes</td>
<td>Italian</td>
<td>0-10</td>
<td>y6 = Yes</td>
</tr>
<tr>
<td>X7</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>None</td>
<td>$</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Burger</td>
<td>0-10</td>
</tr>
<tr>
<td>X8</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Some</td>
<td>$</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Thai</td>
<td>0-10</td>
</tr>
<tr>
<td>X9</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Full</td>
<td>$</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Burger</td>
<td>&gt;60</td>
</tr>
<tr>
<td>X10</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Full</td>
<td>$$$$</td>
<td>No</td>
<td>Yes</td>
<td>Italian</td>
<td>10-30</td>
<td>y10 = No</td>
</tr>
<tr>
<td>X11</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>None</td>
<td>$</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Thai</td>
<td>0-10</td>
</tr>
<tr>
<td>X12</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Full</td>
<td>$</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Burger</td>
<td>30-60</td>
</tr>
</tbody>
</table>

Data

ID3

Decision Tree

S. Russel & P. Norvig. Artificial Intelligence - A Modern Approach
## ID3 - Training Data

<table>
<thead>
<tr>
<th>Example</th>
<th>$Alt$</th>
<th>$Bar$</th>
<th>$Fri$</th>
<th>$Hun$</th>
<th>$Pat$</th>
<th>$Price$</th>
<th>$Rain$</th>
<th>$Res$</th>
<th>$Type$</th>
<th>$Est$</th>
<th>Goal</th>
<th>$Will\ Wait$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Some</td>
<td>$$$</td>
<td>No</td>
<td>Yes</td>
<td>French</td>
<td>0–10</td>
<td>$y_1$ = Yes</td>
<td></td>
</tr>
<tr>
<td>$x_2$</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Full</td>
<td>$</td>
<td>No</td>
<td>No</td>
<td>Thai</td>
<td>30–60</td>
<td>$y_2$ = No</td>
<td></td>
</tr>
<tr>
<td>$x_3$</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Some</td>
<td>$</td>
<td>No</td>
<td>No</td>
<td>Burger</td>
<td>0–10</td>
<td>$y_3$ = Yes</td>
<td></td>
</tr>
<tr>
<td>$x_4$</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Full</td>
<td>$</td>
<td>Yes</td>
<td>No</td>
<td>Thai</td>
<td>10–30</td>
<td>$y_4$ = Yes</td>
<td></td>
</tr>
<tr>
<td>$x_5$</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Full</td>
<td>$$$</td>
<td>No</td>
<td>Yes</td>
<td>French</td>
<td>&gt;60</td>
<td>$y_5$ = No</td>
<td></td>
</tr>
<tr>
<td>$x_6$</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Some</td>
<td>$$</td>
<td>Yes</td>
<td>Yes</td>
<td>Italian</td>
<td>0–10</td>
<td>$y_6$ = Yes</td>
<td></td>
</tr>
<tr>
<td>$x_7$</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>None</td>
<td>$</td>
<td>Yes</td>
<td>No</td>
<td>Burger</td>
<td>0–10</td>
<td>$y_7$ = No</td>
<td></td>
</tr>
<tr>
<td>$x_8$</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Some</td>
<td>$$</td>
<td>Yes</td>
<td>Yes</td>
<td>Thai</td>
<td>0–10</td>
<td>$y_8$ = Yes</td>
<td></td>
</tr>
<tr>
<td>$x_9$</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Full</td>
<td>$</td>
<td>Yes</td>
<td>No</td>
<td>Burger</td>
<td>&gt;60</td>
<td>$y_9$ = No</td>
<td></td>
</tr>
<tr>
<td>$x_{10}$</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Full</td>
<td>$$$</td>
<td>No</td>
<td>Yes</td>
<td>Italian</td>
<td>10–30</td>
<td>$y_{10}$ = No</td>
<td></td>
</tr>
<tr>
<td>$x_{11}$</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>None</td>
<td>$</td>
<td>No</td>
<td>No</td>
<td>Thai</td>
<td>0–10</td>
<td>$y_{11}$ = No</td>
<td></td>
</tr>
<tr>
<td>$x_{12}$</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Full</td>
<td>$</td>
<td>No</td>
<td>No</td>
<td>Burger</td>
<td>30–60</td>
<td>$y_{12}$ = Yes</td>
<td></td>
</tr>
</tbody>
</table>
ID3

- **Strategy**
  - Find feature that “splits” training data the best
    - splits data into Yes inputs and No inputs as much as possible
    - New sets are smaller
  - Recurse on each set

---

S. Russel & P. Norvig. Artificial Intelligence - A Modern Approach
ID3 - Feature Importance

- How do we find a feature to split on?
  - Choose most “important” feature
  - Importance of feature measured using its entropy

- Entropy
  - Quantifies uncertainty of a probability distribution
  - Probability of rain in Sahara desert has low entropy
  - Probability of rain in Providence has higher entropy
  - Probability of rain Seattle?
ID3 - Feature Importance

- Measure entropy of dataset
  - View ratios of Yes's and No's as their probability
  - Compute entropy of that distribution
- Measure importance of a feature
  - Use feature to split inputs into classes
  - For each class
    - View ratios of Yes's and No's as their probability
    - Compute entropy of that probability distribution
    - Combine entropies of classes
  - Compute gain of feature
    - Entropy of data minus combined class entropies
ID3 - Feature Importance

S. Russel & P. Norvig. Artificial Intelligence - A Modern Approach
ID3 - Feature Importance

- Gain of feature
  - data entropy - combined entropy
- Choose feature with largest gain and recurse

S. Russel & P. Norvig. Artificial Intelligence - A Modern Approach
S. Russel & P. Norvig. Artificial Intelligence - A Modern Approach
Outline

- Motivation
- Supervised learning
- Decision Trees
- Algorithmic Bias
ML Applications

- ML algorithms are different than classical algorithms
  - ML algorithms behavior/decisions depend on training data
  - Classical algorithms behavior does not
- ML algorithms are being embedded everywhere
- Traditional applications
  - spell checking
  - ad targeting
  - recommendations
  - speech & handwriting recognition
  - computer vision
ML Applications

- Newer applications
  - News feeds
  - Self-driving cars
  - Risk assessment
- These have serious impact on society
  - should news be tailored to your political beliefs?
  - should car save driver or pedestrian?
  - should we deny freedoms based on risk assessments?
Risk Assessment

- ML algorithms that predict risk
  - Insurance premiums
- Credit
- Crime
  - Recidivism
- Bonds
- Sentencing
Criminal Risk Assessment

- ML in criminal justice system
  - better predict who will commit new crimes
- Purported goals
  - a fairer system
  - less people in jail
  - for less time
Criminal Risk Assessment Tools

- COMPAS by Northpointe predicts
  - Risk of new violent crime
  - Risk of general recidivism
  - Pretrial risk (failure to appear)
- Public Safety Assessment by Arnold Foundation
Criminal Risk Assessment Tools

- Used in
  - Arizona, Colorado, Delaware, Kentucky, Louisiana, Maryland, New York, Oklahoma, Virginia, Washington, Wisconsin, …

- Often adopted without independent study of effectiveness

- Example
  - New York started using tool in 2001
  - Studied it only in 2012
ProPublica Study

- ProPublica conducted a study of COMPAS
  - 7000 arrests in Broward County, FL
  - between 2013 and 2014
- Found scores unreliable for violent crimes
  - 20% of people labeled high risk committed new violent crimes
- Predictions improved for all crimes
  - with misdemeanors included
  - 61% of people labeled high risk committed new crimes
ProPublica Study

- Found significant racial disparities
- Labeled high risk but didn’t re-offend
  - African-americans: 44.9%
  - Whites: 23.5%
- Labeled low risk but did re-offend
  - African-american: 28%
  - White: 47.7%
- Study accounted for
  - Criminal history, age and gender
Algorithms

- Algorithms are profoundly impacting our lives
- You now know how to
  - design algorithms
  - analyze algorithms
  - implement algorithms
- Don’t forget that…
  - …your algorithms will impact people
Post CS16

- Very solid foundation in CS
- You can start to learn about all areas of CS
  - from most applied to most theoretical
  - Advanced Algorithms, Randomized Algorithms
  - Operating systems, Networking, Distributed Systems
  - Programming Languages
  - Graphics, Animation & Visualization
  - Human-Computer Interaction
  - Machine Learning, Computer Vision, Natural Language Processing
  - AI & Robotics
  - Security, Cryptography & Privacy
Announcements

- No Lecture next 2 weeks
- Review sessions
  - May 9th
  - May 11th
Thank You
UTAs

- Alejandro Molina
- Brian Lee
- Chantal Toupin
- Data Storer
- Elaine Jiang
- Hannah Tipperman
- Justine Breuch
- Jewel Brown
- Jonathan Chemburkar
- June Ge
- Katie Chu
- Kaila Jeter
- Marcus Yeo
- Laura Blackstone
- Lauren Ho
- Michael Chen
- Montana Fowler
- Nia Sanders
- Rachel Teller
- Reca Sarfati
- Sophia Hsiao
- Sarah Kim
- Sophie Saskin
- Tristin Falk-LeFay
- Lucy Huang
- Michael Chen
- Zach Kirschenbaum
HTAs

- Devanshi Nishar
- Kei Nakagawa
- Surbhi Madan
- Wesley Herts
References

- *Artificial Intelligence - A Modern Approach* by S. Russel & P. Norvig
- *Machine Bias* by ProPublica
- *New York State COMPAS-Probation Risk and Need Assessment Study*
- *Evaluating the Predictive Validity of the COMPAS Risk and Needs Assessment System* by Northpointe
References

- Slide #2
  - Toothless from *How to Train Your Dragon*