Visualizing Models and Methods for Decision-makers

Michael Freeman
University of Washington
@mf_viz
a plea...
{Where does data come from?}
These sources can be pretty important.
they can be biased
there may be assumptions
Help keep your streets smooth

RECORD A TRIP

MY TRIPS
Decision-makers should understand these sources.
What happens here?
“Above all else, show the data”
These actions can be pretty important.
we make mistakes
Bank of America Finds a Mistake: $4 Billion Less Capital

By PETER EAVIS and MICHAEL CORKERY

April 28, 2014 9:18 AM

92 Comments

Following the accounting error, Brian Moynihan, Bank of America's chief, will have to again rebuild trust in the bank. Drew Angerer/Getty Images
we make assumptions
How we amplify privilege with supervised machine learning

Michael Williams // @willi RMSMJ

Attempt to identify negative sentiment

Train a tool that identifies harsh language

People using severe language get attention

{assumes men and women use strident language at the same rate}
Decision-makers should understand these processes.
What do we talk about?
{user-centered design}
User-centered design

Who is your audience?
What is their question?
What is their level of familiarity?
What is the impact of their decisions?
{visualize}
Choose the graphical encodings…

… that are best visually decoded.

Can we do this with methods and models?
which values do we encode?
Visually encoding methods

Express the **structure** of the data

Encode the **process** employed

Display any **impact** on the data
expressing structure
Global Health Funding

Source
(Where money comes from)

Governments

Gates Foundation
<table>
<thead>
<tr>
<th>Source</th>
<th>Channel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Governments</td>
<td>United Nations</td>
</tr>
<tr>
<td>Gates foundation</td>
<td>Global Fund</td>
</tr>
<tr>
<td></td>
<td>Governments</td>
</tr>
</tbody>
</table>
## Global Health Funding

<table>
<thead>
<tr>
<th>Source</th>
<th>Channel</th>
<th>Recipient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Governments</td>
<td>United Nations</td>
<td>Health focus area</td>
</tr>
<tr>
<td>Gates foundation</td>
<td>Global Fund</td>
<td>Region</td>
</tr>
<tr>
<td>Governments</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
{encoding process}
Linear regression

Find a function that describes the relationship between two+ variables

Done by minimizing the squared distance between data and function
Linear regression

hand size = height

credit: setosa.io
Markov chains

*Used in simulations as a random process for state transitions*

*Probability distribution of the next state depends only on the current state*
Markov chains

- Bull market
- Bear market
- Stagnant market

Transition probabilities:
- From Bull to Bear: 0.02
- From Bear to Bull: 0.3
- From Bull to Stagnant: 0.005
- From Stagnant to Bull: 0.02
- From Bear to Stagnant: 0.4
- From Stagnant to Bear: 0.2
Machine learning

{let’s just look at the visual}
Machine learning

credit: r2d3.us
{displaying impact}
function shuffle(array) {
    var n = array.length, t, i;
    while (n) {
        i = Math.random() * n-- | 0; // 0 ≤ i < n
        t = array[n];
        array[n] = array[i];
        array[i] = t;
    }
    return array;
}
function shuffle(array) {
    var n = array.length, t, i;
    while (n) {
        i = Math.random() * n-- | 0; // 0 ≤ i < n
        t = array[n];
        array[n] = array[i];
        array[i] = t;
    }
    return array;
}

// DON'T DO THIS!
function shuffle(array) {
    return array.sort(function(a, b) {
        return Math.random() - .5; // ಥ_ಥ
    });
}
This is good!
This is bad!
This is **firefox**!
so what should you do?
Understand
Users

Express
Structure

Encode
Process

Display
Impact
Credits

Financing Global Health (IHME)
   http://vizhub.healthdata.org/fgh/

Linear Regression (setosa.io)
   http://setosa.io/ev/ordinary-least-squares-regression/

Markov Chains (setosa.io)
   http://setosa.io/ev/markov-chains/

Machine Learning (r2d3)
   http://www.r2d3.us/visual-intro-to-machine-learning-part-1/

Shuffling (Mike Bostock)
   http://bost.ocks.org/mike/algorithms/#shuffling
In order to understand data...

... you need to understand what you did to it.
Thank you

Michael Freeman
mikefree@uw.edu
mfviz.com
@mf_viz
mfviz.com/strata-2015