Machine Learning and Democracy: Some Problems in Collective Decision-Making

Sanjeev R. Kulkarni
Professor, Department of Electrical Engineering
Dean of the Graduate School
Princeton University

Joint work with Tiance Wang, Paul Cuff
Outline

• Science of information meets the liberal arts
• Artificial intelligence and machine learning
• Some problems in collective decision-making
  • Recommender systems
  • Aggregating probability judgments
  • Voting and social choice
A Broad View of the Science of Information

Information-based Paradigm for Designing Systems

Systems may be complex, dynamic, distributed
Some Information-Processing Tasks

- Sensing
- Sampling, Quantization (& D/A)
- Filtering
- Storage and Representation
- Search and retrieval
- Compression
- General purpose computation
- Communication, Data Transmission
- Error Detection/Correction
- Cryptography
- Digital Rights Management
- Learning and Inference
- Control
- Actuation
Where, Why, How to Meet the Liberal Arts

• Connections to many areas: mathematics, psychology, philosophy, economics, politics, public policy, physics, biology, linguistics, art, classics, etc.

• It’s all around us and it’s interesting

• A liberal arts education in the 21st century should include some basic understanding of technology (including information technology)

• Teaching and research at the intersection:
  • Inherently blended fields
  • Address technology in its broader societal context
  • Similar problems and sharing/accessibility of ideas and tools

Then

Now
Artificial Intelligence and Machine Learning
• (John McCarthy, Stanford University) The science and engineering of making intelligent machines, especially intelligent computer programs

• But what’s intelligence? No satisfactory general definition
Some Fundamental Problems Within AI

Learning

Natural Language Processing

Deduction, Reasoning, Problem Solving

Motion and Manipulation

Knowledge Representation
Planning
Creativity

Perception (vision, hearing, etc.)
The Modern Beginning

Alan Turing (1912-1954), founder of modern computer science

Proposed “Turing Test” as a test for a machine to demonstrate intelligence. (1950)

Judge tries to distinguish human from machine.

Dartmouth Conference, Summer of 1956

“The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves.”
• Some early successes in a variety of areas (natural language processing, search, game playing)

• Wild enthusiasm and optimism

• Bold predictions:
  • 1965, H. A. Simon: "machines will be capable, within twenty years, of doing any work a [hu]man can do."
  • 1967, Marvin Minsky: "Within a generation … the problem of creating 'artificial intelligence' will substantially be solved."
  • 1970, Marvin Minsky: "In from three to eight years we will have a machine with the general intelligence of an average human being."

• Then reality started setting in…
The spirit is willing but the flesh is weak.

  The alcohol is arranged but the meat is weak.

- (2010) 這種精神是願意，但心有餘而力不足。
  This spirit is willing, but beyond their grasp.

- (Urban legend?) Водка хороша, но мясо гнилое:
  The vodka is good but the meat is rotten.

Out of sight, out of mind.

- Invisible insanity.
- Blind idiot.

*AI and Machine Learning*
Winter, Boom, Winter

- **AI Winter 1 (mid 1970’s to 1980)**
  - Difficulty of tasks recognized.
  - Disillusionment set in.
  - Funding evaporated.

- **Boom (1980’s)**
  - Expert systems, knowledge-based systems, connectionist revival.
  - Some successes renewed optimism.
  - Funding returned.

- **AI Winter 2 (late 1980’s-early 1990’s)**
  - Systems expensive, brittle, difficult to maintain.
  - Special-purpose hardware falls prey to advances in desktops.
  - Failed once again to meet grandiose expectations.
Another Boom: Machine Learning (1990’s – present)

- Develop methods to automatically draw inference from different types of data

- Applications to recognition of images (faces, objects, etc.), speech, handwriting, medical diagnosis, spam, fraud, advertising, bioinformatics, etc.

- How can we design effective algorithms?

- Can we understand the fundamental limits? What can be learned? What can’t? Why?
Another Boom: Machine Learning (1990’s – present)

- Huge advances in computing power and other areas
- Using more tools from mathematics and other fields
- Transition from too little data to too much data
- Some notable successes

- IBM Deep Blue vs. Kasparov, 1997
  - 6 games: K, DB, draw, draw, draw, DB

- DARPA Grand Challenge, 2005
  - 131 miles of autonomous driving in desert

- IBM’s Watson, vs. Ken Jennings and Brad Rutter in Jeopardy, Feb 2011
  - Watson wins

- Tremendous progress over past 50 years – many algorithms, theoretical results, applications. (Though still nowhere near original expectations of AI.)
Connections to Many Other Fields and Many Fundamental Questions

- Policy, Privacy, Security
  - Massive amount of data (internet traffic, cell phones, geolocation, social networking, cameras, databases)
  - What can we do with it? How should we use it?
  - Who owns it? How safe is it? Rights, rules, responsibilities?
  - What can be inferred? Anonymization?

- Mathematics, optimization, statistics

- Neuroscience, cognitive science, psychology

- Art, literature, classics, politics

- Philosophy
  - Induction, simplicity, Occam’s razor
  - Mind, consciousness, awareness, intent, feeling
  - Ethical implications?
A Current Challenge: Collective Decision-Making
A Current Challenge: Collective Decision-Making

- Collection of agents
- Each with some “data”
  - Belief
  - Preference
  - Other information

Smart Dust  Mobile Sensors

- Data can be distributed, dynamic, heterogeneous
- Resource constraints
- Wish to aggregate the data (make a collective decision)
Recommender Systems

- With internet, huge number of possible items.
- Would like to find some new items that we believe a user would rank highly.
- Wide range of applications: product recommendations, movies, news, etc.
Recommender Systems: Methods and Issues

<table>
<thead>
<tr>
<th></th>
<th>i1</th>
<th>i2</th>
<th>i3</th>
<th>i4</th>
<th>i5</th>
<th>i6</th>
<th>i7</th>
<th>i8</th>
</tr>
</thead>
<tbody>
<tr>
<td>u1</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>u2</td>
<td></td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>u3</td>
<td>3</td>
<td>?</td>
<td></td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>u4</td>
<td>5</td>
<td>2</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>u5</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

- **Methods**: content-based filtering and collaborative filtering
- **Issues**:
  - Scalability (huge number of items and users)
  - Sparsity (very little data)
  - Cold start (how to recommend for new users)
  - Attacks (some users may wish to game the system)
Aggregating Probability Forecasts

• Get probability forecasts on a number of interrelated, complex events.

• Forecasts likely disagree and may not even be internally compatible (humans are notoriously inconsistent).

• Would like to aggregate probability estimates.

• Wide range of applications: geopolitical forecasting, economics, finance, sensor networks.
Voting and Social Choice

- Majority vote is obvious way to make choice between two alternatives
- With more alternatives, things get more complicated.
- What are good methods for collective decision-making?
Some Problems With Majority Voting
Discursive Dilemma/Doctrinal Paradox

<table>
<thead>
<tr>
<th></th>
<th>Judge 1</th>
<th>Judge 2</th>
<th>Judge 3</th>
<th>Majority</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>T</td>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>p implies q</td>
<td>T</td>
<td>T</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>q</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
</tbody>
</table>

Majority vote leads to inconsistent outcomes!
### Some Problems With Majority Voting

#### Discursive Dilemma/Doctrinal Paradox

<table>
<thead>
<tr>
<th></th>
<th>20%</th>
<th>40%</th>
<th>40%</th>
<th>Majority</th>
</tr>
</thead>
<tbody>
<tr>
<td>New road helpful?</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y (60%)</td>
</tr>
<tr>
<td>Funds available?</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y (60%)</td>
</tr>
<tr>
<td>Build new road?</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N (80%)</td>
</tr>
</tbody>
</table>

Outcome depends on how agenda is structured!
<table>
<thead>
<tr>
<th></th>
<th>1^{st} choice</th>
<th>2^{nd}</th>
<th>3^{rd}</th>
<th>4^{th}</th>
</tr>
</thead>
<tbody>
<tr>
<td>1^{st} choice</td>
<td>Clinton</td>
<td>Trump</td>
<td>Johnson</td>
<td>Stein</td>
</tr>
<tr>
<td>2^{nd}</td>
<td>Stein</td>
<td>Johnson</td>
<td>Stein</td>
<td>Trump</td>
</tr>
<tr>
<td>3^{rd}</td>
<td>Johnson</td>
<td>Stein</td>
<td>Trump</td>
<td>Clinton</td>
</tr>
<tr>
<td>4^{th}</td>
<td>Trump</td>
<td>Clinton</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

How should we decide which candidate wins?
Plurality Voting

<table>
<thead>
<tr>
<th>1&lt;sup&gt;st&lt;/sup&gt; choice</th>
<th>Clinton</th>
<th>Trump</th>
<th>Johnson</th>
</tr>
</thead>
<tbody>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt;</td>
<td>Stein</td>
<td>Johnson</td>
<td>Stein</td>
</tr>
<tr>
<td>3&lt;sup&gt;rd&lt;/sup&gt;</td>
<td>Johnson</td>
<td>Stein</td>
<td>Trump</td>
</tr>
<tr>
<td>4&lt;sup&gt;th&lt;/sup&gt;</td>
<td>Trump</td>
<td>Clinton</td>
<td>Clinton</td>
</tr>
</tbody>
</table>

- Plurality: pick candidate with most 1<sup>st</sup> choice votes
- **Clinton wins!**
### Instant Run-off

<table>
<thead>
<tr>
<th></th>
<th>45%</th>
<th>40%</th>
<th>15%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1st choice</strong></td>
<td>Clinton</td>
<td>Trump</td>
<td>Johnson</td>
</tr>
<tr>
<td><strong>2nd</strong></td>
<td>Stein</td>
<td>Johnson</td>
<td>Stein</td>
</tr>
<tr>
<td><strong>3rd</strong></td>
<td>Johnson</td>
<td>Stein</td>
<td>Trump</td>
</tr>
<tr>
<td><strong>4th</strong></td>
<td>Trump</td>
<td>Clinton</td>
<td>Clinton</td>
</tr>
</tbody>
</table>

- Instant Run-off: successively eliminate candidate with fewest 1st choice votes
- Delete Stein, then Johnson, then Clinton
- **Trump wins!**
### Borda Count

<table>
<thead>
<tr>
<th></th>
<th>1st (4 pts)</th>
<th>2nd (3 pts)</th>
<th>3rd (2 pts)</th>
<th>4th (1 pt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>45%</td>
<td>Clinton</td>
<td>Stein</td>
<td>Johnson</td>
<td>Trump</td>
</tr>
<tr>
<td>40%</td>
<td>Trump</td>
<td>Johnson</td>
<td>Stein</td>
<td>Clinton</td>
</tr>
<tr>
<td>15%</td>
<td>Johnson</td>
<td>Stein</td>
<td>Trump</td>
<td>Clinton</td>
</tr>
</tbody>
</table>

- **Borda count**: 4 points for 1\textsuperscript{st}, 3 points for 2\textsuperscript{nd}, etc.
- Johnson \( (0.45)(2) + (0.40)(3) + (0.15)(4) = 2.7 \)
- Clinton 2.35, Trump 2.35, Stein 2.6
- **Johnson wins!** (Although any of the others could win with different assignment of points.)
What method should we use?

Two separate issues:
- What information to collect from individuals
- How to combine these into a social (democratic) choice

Some possibilities for information to collect:
- Top choice
- Ranking
- Acceptable choices
- Score (say from 0 to 100)

For combining, must have/would like:
- Voters and candidates treated equally
- Unanimity: If A>B for all voters, then society decides A>B
- Transitive social preferences: A>B and B>C implies A>C
- Non-manipulable
- Independence of irrelevant alternatives
Independence of Irrelevant Alternatives

<table>
<thead>
<tr>
<th></th>
<th>45%</th>
<th>40%</th>
<th>15%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt;</td>
<td>Clinton</td>
<td>Trump</td>
<td>Johnson</td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt;</td>
<td>Stein</td>
<td>Johnson</td>
<td>Stein</td>
</tr>
<tr>
<td>3&lt;sup&gt;rd&lt;/sup&gt;</td>
<td>Johnson</td>
<td>Stein</td>
<td>Trump</td>
</tr>
<tr>
<td>4&lt;sup&gt;th&lt;/sup&gt;</td>
<td>Trump</td>
<td>Clinton</td>
<td>Clinton</td>
</tr>
</tbody>
</table>

- Suppose we use plurality voting
- Recall Clinton wins

- What if Johnson and Stein weren’t running? Then Trump wins.
- Adding Johnson and Stein changes the outcome. Should adding irrelevant alternatives change things?
- Seems not for individuals (e.g., Sidney Morgenbesser story). Maybe not for groups either.
- Plurality is very poor regarding IIA.
- Arrow’s Impossibility Theorem: The only procedure that satisfies unanimity, transitivity, and IIA is a dictatorship!
Suppose we use plurality voting.

Suppose voters know preferences (say from polls).

15% incentivized to vote strategically so Trump beats Clinton.

Are there reasonable methods that are immune to strategic voting?

No! Gibbard-Satterthwaite Theorem: Every deterministic rule with 3 or more candidates satisfies at least one of the following:

- The rule is dictatorial.
- There is some candidate who can never win.
- The rule is susceptible to strategic voting.
Voting Profile, Boundaries, and Strategic Voting

- Consider ranked voting systems that are anonymous (treat all voters equally).
- With M candidates, there are M! possible rankings.
- Voting profile can be represented by the number (or fraction) of voters who prefer each of these rankings. I.e., the profile is a point in the M! simplex.
- A deterministic voting system is a mapping from this simplex to the set of candidates.
- Equivalently, it’s a partition of the simplex into M labeled regions.

- The boundaries of the regions are manipulable since a single voter can change the outcome.
- Boundaries can be either strategic or non-strategic.
- Gibbard-Satterthwaite shows that for any deterministic voting system (mapping), there exists at least one profile that would lead to strategic voting.
Are Some Methods Less Susceptible to Strategic Voting?

- Yes! Condorcet methods are categorically more resistant to strategic voting.
- Condorcet Method: Selects the candidate (if there is one) that wins in all pairwise matchups.
- There isn’t always a Condorcet winner. Different Condorcet methods give different outcomes in this case.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Candidate</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Clinton</td>
<td>45%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Stein</td>
<td>40%</td>
<td>15%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Johnson</td>
<td></td>
<td>33%</td>
<td>15%</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Trump</td>
<td></td>
<td></td>
<td>44%</td>
<td>20%</td>
</tr>
</tbody>
</table>

- Johnson wins all pairwise matchups
- Johnson is the Condorcet winner
• Strategic voting requires some knowledge. If a voter knows nothing about other voters’ choices, there’s no incentive (information with which) to strategize.

• If voters know the profile perfectly, they will strategize.

• Assume the actual profile is random centered around some mean profile p.

• Assume voters know p (e.g., through polls), but don’t know the actual profile.

• A voter can find nearest boundary and vote strategically or not, depending on the boundary.

• I.e., if the other side of the boundary leads to a better outcome, then it’s a strategic boundary.
Condorcet Methods and Strategic Voting

- Condorcet methods have some non-strategic boundaries. Other common methods don’t.
- Can show that probability (under certain models) of strategic voting significantly smaller than Plurality or Borda Count.
- Some data we gathered also shows strategic voting less likely for Condorcet methods.

Condorcet | Borda | Plurality
Condorcet Methods: More Good News

- A Condorcet winner seems likely to exist in practice.

- Condorcet methods uniquely satisfy IIA whenever possible.

- In some dynamic settings (with voters sequentially allowed to change their votes based on feedback), Borda count converges to the Condorcet winner (if there is one).
Dynamic Setting

• Assume every voter has an honest preference, acts independently, and is strategic.

• Uniformly pick a random voter.

• Voter observes noisy version of current profile and recasts vote based on nearest boundary.

• Borda count converges to Condorcet winner if there is one. Doesn’t converge if there’s no Condorcet winner.

• Plurality converges to Instant Run-off voting.
Condorcet Methods: Some Bad News

- Need rank information from voters
- Need method to select a winner if there is no Condorcet winner
  - E.g., Kemeny-Young: Find the ranking that minimizes the sum of the Kendall-Tau distance to each voter’s preference. Select a winner based on this ranking.
- Harder to compute
- Harder to understand
A Plug for Approval Voting

• Approval Voting: Vote for as many candidates as you like (i.e., that you approve of). Winner is candidate who receives most approval votes.

• Many advantages
  • Easy to understand voting and aggregating process.
  • Easy to implement at ballot box.
  • Fewer spoiled ballots.
  • Easy to compute/tally votes.
  • Gives proper due to minority/third-party candidates (more accurate measure of support).
  • Less susceptible to strategic voting. Satisfies the “Favorite Betrayal Criterion”.

• But... some disadvantages
  • Candidate who is first choice of majority can lose
  • Can lead to candidate many find just “okay”
So Where Do We Go From Here?

- No democratic process (information obtained and aggregation method) dominates all others.
- How should we decide what method to use?

- Fairness of process and quality of outcome?
- Simplicity versus fidelity to will of the people?
- Tradeoff between different desirable properties.
- Role of technology
  - Aggregation algorithms and analysis
  - Agenda setting
  - Direct democracy versus representative democracy
• Many interesting areas of intersection between science of information and the liberal arts.

• Machine learning (and collective decision-making in particular) is one such area.

• Technological advances have made collective decision-making an important and practical problem in many applications.

• Related to fundamental questions of democracy.

• Democratic process is elusive. Many ways to assess the will of the people, and no method is perfect.

• But some methods are better than others.

• Much is understood but much more to be done.

• Technology and tools from the science of information and decision-making are sure to play an increasing role.
Thank You!