Mining Data Related to Children: Ethical Challenges

Daya C. Wimalasuriya, University of Moratuwa, Sri Lanka
Dejing Dou, University of Oregon, USA
Dilini Wimalasuriya, Northshore College of Business and Technology, Sri Lanka

Introduction

- *Data mining your children*, Stephanie Simon, Politico News Portal, May 2014

  "Data is the real asset" – attributed to Sal Khan, founder of Khan Academy

  "It's free. But users do pay a price: In effect, they trade their data for the tutoring"

- Shows the perils encountered by organization mining data related to children.

Why is this an issue?

- Children are vulnerable; the society should protect them.

- Laws that apply for data related to children especially in an educational setting.

- History of ethically questionable marketing practices towards children: "*they aren't children so much as what I like to call 'evolving customers'* "

This is not the whole truth

- The Khan Academy, supported by data mining, has helped thousands of students.

- The same is true of Google Apps for Education.

- What is necessary is to address the legitimate concerns regarding mining data related to children.

Background

- Laws

- Marketing for Children

- Technology

Laws

- Children's Online Privacy Protection Act of 1998 (COPPA) is the primary legislation in the US that applies to children's Internet usage.

- Applies for children under the age of 13.

- Imposes requirements on websites dealing with children under the age of 13.

  e.g, obtaining verifiable parental consent

- The outcome is that web applications don't deal with children under the age of 13 at all.

- Serves as a warning to the data mining community.

More Laws

- The Family Education Rights and Privacy Act of 1974 (FERPA) applies to student education records held by an organization.

- The education records can be released only with the consent of the student (if she is above 18) or parents.

Marketing for Children

- Main stages of a child evolving as a customer (McNeal)

  - From age 1: Accompanying parents and observing.

  - From age 2: Accompanying parents and requesting.

  - From age 3: Accompanying parents and selecting with permission

  - From age 4: Accompanying parents and making independent purchasing decisions.

  - From age 5: Going to store and making independent purchases.

- Purchasing decisions affected by children (Beder)

  - Made by children by themselves

  - Parental decisions influenced by children

  - Purchases made once they grow up

Marketing for Children Efficiently

- American Psychological Association (APA) has appointed a task force (2000) to
  - Study the effect of advertising on children
  - Investigate the issue of psychologists collaborating with marketers
- Recommendations
  - Several recommendations for public policy such as restricting advertising primarily directed towards audiences of young children
  - "undertake efforts to help psychologists weigh the potential ethical challenges involved in professional efforts to more effectively advertise to children, particularly those children who are too young to comprehend the persuasive intent of television commercials"

Technology

- Personalized advertisements is one of the most visible applications based on mining user data.

- Involves building up a user profile and serving ads that are of interest to the user.

- Cookies are used for tracking

- Do-Not-Track initiative has failed to get much traction so far.

- More sophisticated data mining (e.g., association rule mining, clustering) carried by owners of large user data sets (e.g., social networks)

Ethical Concerns

- Acceptability of the practice

- Differential treatment in educational settings

- Transfer of data

- Transfer of mined knowledge

- Personalized ads

- Transition from an adult to a child

Acceptability of the Practice

- Yes
  - because of the benefits it can produce
  - because of antitrust concerns
- Caveats
  - COPPA strongly indicates that the data should only be used for the purpose they were collected for.
  - Special care should be taken.

Differential Treatment in Educational Settings

- Much of the negative coverage is about data on an educational setting.

- The implication is that such usage on such data should be held to a higher standard.

- Mine data only for enhancing students' learning experience.

- The industry is already moving in this direction.

Transfer of Data

- Raises ethical concerns even with anonymization
  - Anonymization techniques can often be broken
  - "Repurposing" of data
  - "Data brokers" building up detailed profiles of people
- Should be considered a high risk endeavor

Transfer of Mined Knowledge

- Safer than transfer of data

- May still raise ethical concerns (e.g., attributes of children consuming a lot of fast food)

- Ethical concerns have to be handled on a case-by-case basis

Personalized Advertisements

- The concerns are on the cases of inappropriate ads not the practice itself.

- "No system is perfect. When we find or are made aware of prohibited ads, we remove them immediately" – Facebook

- Better technology

From Child to an Adult

- A child would have left a considerable digital footprint by the time she turns 18.

- Is it ok to combine such data with data they produce as adults and use as we see fit?

- Methods of hard or soft separation of such data should be investigated.

Conclusion

- The data mining industry and community should pay special attention to ethical issues ethical issues, especially when dealing with data related to people.

- Ignoring such concerns raises considerable dangers for the industry.

Thank You.

Q & A