GONZAGA
UNIVERSITY

AI, Information Technologies, and Fraud Detection: Ethical Issues

September 16, 2020

Why Is It Important?

"Data analytics and AI are the future of internal audit and fraud investigation"

--Smart Business Network, Inc.

www.sbnonline.com

Today's Objectives

- What is AI? What is its value?
- Data Mining & Data Analytics
- Information Technologies & Privacy
- AI and the Presumption of Innocence
- What Data Mining & Analytics Say About You

What Is AI and What Is Its Value?

What Is Artificial Intelligence?

"Intelligence demonstrated by machines in contrast to the natural intelligence displayed by humans"

-- *Wikipedia*

Helpful???

"Intelligence" simply means human reasoning as it exists

Just machines doing what we do

What Is Artificial Intelligence?

Alan Turing (1950's) -- The Turing Test

In order for a machine to be intelligent, a human being should be unable to distinguish between the machine and a human being based just on the replies given to questions

What Is Artificial Intelligence?

"Algorithms enabled by constraints, exposed by representations that support models targeted at loops that tie thinking, perception and action together"

-- Patrick Winston, MIT

More helpful???

Attempt to explain the "intelligence" part

AI – General View

"Machines that respond to stimulation consistent with traditional responses from humans, given the human capacity for contemplation, judgment, and intention."

-- John Allen & Darrel West, "How Artificial Intelligence is Changing the World," Brookings Institute (www.brookings.edu)

Breaks out "intelligence" into three components

AI – Replicative Model

Purpose of AI is to replicate, mimic human reasoning, but faster

Examples: Understanding human speech

Playing strategic games like chess

Self-driving cars

All things we can do, but AI does it “better”

AI – Perfectionist Model

Retains some of the pieces of the replicative model

- Basic logic
- Cost-benefit reasoning
- Pattern recognition

Remove the cognitive biases that plague human reasoning –"perfect reasoning"

-- e.g. overvaluing losses and implicit biases

Machine Learning

Building a mathematical model based on “training data” in order to make predictions or decisions without being explicitly programmed to do so – algorithm “learns” to recognize patterns -- e.g. email filtering

Supervised – algorithm is trained using labelled pairs of inputs and outputs, e.g. emails with subject line of “X” are spam

Unsupervised – algorithm looks for patterns in data sets without pre-labelled categories to put patterns in

AI Preempting Bad Behavior

"JPMorgan Algorithm Knows You're a Rogue Employee Before You Do"

-- *Bloomberg*, April 2015

AI Preempting Bad Behavior

"Tracking Traders' Emotions"

-- *Bloomberg*, September 2016

AI Catching Fraud

"Using Artificial Intelligence, Visa is Combatting Fraud at Nearly the Speed of Light"

-- *PaymentsJournal*, June 17, 2019

Data Mining & Data Analytics

Data Mining

Accessing databases containing large amounts of historical data – doesn't produce information per se, unless combined with some data analytics tool

Google Analytics

Retailers segment customers into Recency, Frequency, Monetary' (RFM) groups; target marketing to those different groups.

Police depts. Deciding where to deploy police manpower

Data Analytics

Broader category than AI, though AI is a subset of advanced analytics – AI not used only for analytics uses (e.g. autonomous vehicles)

Examples

Health Care – analyze large data sets of treatment histories to target treatment for particular patient
Netflix – uses data analytics for targeted advertising
Large banks – uses data analytics as part of risk management

Data Analytics in Auditing

Hewlett Packard

Concern with the volume and frequency of manual journal entries

Put in place a continuous auditing and monitoring system that used data analytics to identify trends in the movement in accounts, spikes of activity in accounts, and the identity of individuals posting the manual entries

Value of Data Analytics for Auditing

Allows auditors to access much larger sets of transactions rather than just sampling from that set

More accurate “picture” of the firm

Greater efficiency due to speed of data gathering and organization

Frees auditors from “mechanical tasks,” allowing more attention to the “auditing” part of auditing (more focus on risk assessment)

Other Information Technologies

Asset management software, data integration, gamification techniques, & information security

Simplicable Website
"77 Types of Information Technologies

https://simplicable.com/new/types-of-information-technology

Issues with AI, Data Analytics, and Data Mining

Time Lag Problem

Supervised AI learning and the running of advanced analytics tools require databases, which take time to assemble before they can be used

Typically fraud isn't detected by humans until after the fact

So, fraud and other auditing irregularities will continue to occur as training data is assembled and updated

Problem of False Positives

Want AI or analytics tool to pick out ***only*** the data points relevant to fraud detection or other auditing concerns not irrelevant or outdated data points

If not, potential for false positives – one estimate that only 1 out of every 100 transactions flagged as fraudulent

Creates more work for human auditors and fraud managers

Problem of False Negatives

Want AI or analytics tool to detect ***all*** patterns indicative of the presence of fraud or auditing concerns

If not, get false negatives and fraud or bad behavior goes undetected

Success of AI and analytics tools only as good as the data used or accessed

Questions About Data Access

Data Access

Select the data/information algorithm has access to in supervised training – i.e. “trainers” select the information, make decisions about what and how information to input

Unsupervised training, algorithm is allowed access to a volume of data to search for patterns – a “good trainer” will want as much data as possible or allow algorithm “freedom” to search within databases

What data to use? How much to use?

Ethical Issues

Privacy

Privacy seems to involve access to information about ourselves

"A person has privacy to the extent that others have limited access to information about him, limited access to the intimacies of his life, or limited access to his thoughts or his body" – Schoeman, "Privacy: Philosophical Dimensions of the Literature," 1984

Sphere or Zone of Privacy

Often couched in terms of a "right to privacy" which requires more than utilitarian considerations to override

Privacy

AI trainers and data analysts have to make decisions about which data to allow algorithms to access – choices must be made about which facts of cases to input – e.g. transaction location data and type of purchase data

Both types of data are relevant in many cases, but is it a violation of privacy to compile "picture" of our travels and "picture" of our purchasing habits?

How information is gathered and who has access to it

Is it stored safely?

Privacy

Recall the “Tracking Traders’ Emotions” Case

Physiological data, understood as indicative of emotional states, acquired and used by algorithm

Is coming to know a trader’s emotional states a privacy invasion?

Privacy

Some AI fraud systems scour social media for data to create “fraud risk profile”

What does the algorithm tag as fraud relevant information?

-- that I once dated a convicted fraudster?
-- I Google “How to commit fraud?”

Must I be informed when my social media data accessed by third parties?

Privacy

Number of data analytics tools/algorithms do the same thing -- they scour social media or search engine activity for data to create "demographic profiles" and even individual profiles

For example, automated candidate sourcing – use of data analytic algorithms to create candidate pools based on social media presence, among other things

Same privacy concerns here

Presumed Innocence

Our default justice position – guilt must be established beyond some level of doubt

Rate of false positives an ethical concern -- “falsely accused”

Since the digital world is forever, identification as a fraudster may remain in databases somewhere

No Net Benefit

Given that financial institutions have a right to protect themselves against fraud – so, could justify a privacy intrusion

High rate of false positives creates costs to address and has negative impact on firm reputation – Bank Compliance Officer, “Too much risk of innocent people getting caught up in a dragnet”

Costs not balanced by gains use of using AI – false negatives, fraud gets through anyways

Unintended Discrimination

Amazon's attempt at AI recruiting tool

Algorithm would review resumes and provide a ranking of top candidates – supposedly would save time and be bias free – instead, since trained on ten year database of job applicant resumes, vast majority of which were male, AI tool treated being male as desirable characteristic – consequence, resumes with the word "woman" or "women's" in them were rated lower

Unintended Discrimination

Nine states including Washington, algorithm generated risk assessment reports are used in sentencing process

Study of 7,000 cases from Broward County FL showed that

Algorithm wrongly labelled black offenders as future criminals twice the rate of white offenders
White defendants mislabeled as low risk more often than black defenders

Unintended Discrimination

Apple co-founder Steve Wozniak received a new Apple credit card with a credit limit ten times higher than his wife, Janet Hill, despite sharing bank accounts along with the rest of their assets.

Thanks!!
Questions??