

MACHINE LEARNING APPLICATION IN FRAUD DETECTION.

INSTRUCTOR: DR. ZOYA KINSTLER, TA: TAKAYUKI IIDA

Research Paper "A Digital Technology for a Business Enterprise"

Write a 500-word paper, on the topic of "A Digital Technology for a Business Enterprise".

The paper's purpose is to educate the class about a new technology trend and to explain how a business enterprise may take advantage of it.

Who should be your author persona, whose writing voice should you choose? For this assignment, we would like you to adopt a viewpoint of an enterprise IT leader - innovator and architect, with a strong business sense and a concern for application system impacts.

THE "WHAT" QUESTION

Describe your selected technology, what it is, where it's used, who makes it, how and by whom it's being adopted

THE "HOW" QUESTION

How is the thing built, how does its technology work, how different is it from previous technologies?

THE "WHY" QUESTION

Why should a business enterprise invest in this technology? What known problem can it solve?

MAKE A RECOMMENDATION

Summarize your findings and make a recommendation, what an enterprise needs to do for adopting the technology into its business model and its systems platform.

WHERE TO LOOK:

- Research articles on the class website. A reading list is provided on this class website: on the "Reading and Research" page, open the "Technology Trends for Digital Enterprise" module
- Gartner research – access via Harvard CIO website
- TechCrunch, Information Week and CIO Magazine – free online subscriptions
- Academic journals via Harvard online libraries
- White papers by Enterprise IT software and service companies: IBM, Oracle, SAP, Microsoft, Google, Accenture, and cloud-based business application vendors: salesforce.com, workday.com, servicenow.com, and others
- Technology blogs
- Your research from other classes and your work experience.

WRITE YOUR PAPER:

- Paper size not to exceed 500 words (roughly 1 page), the 2nd page can be used only for pictures and diagrams. Format: MS Word, not pdf; font size from 10 to 12 points; approximate line spacing 1.5
- Please don't forget to include a list of literature references at the end of your paper
- Write well: present your thoughts clearly and logically
- Style to use – "technical-academic" – formal, serious, and as simple as possible.
- In this course we expect students to use precise language, following the writing requirements for graduate level papers at Harvard University. The preferred writing style is "technical-academic", namely: clear, concise, to the point, easy to understand, and rigorously logical. Try to avoid any ambiguities and generalities. Stay away from circuitous "commercial writing" language, replete with empty rhetoric, commonly found in industry literature.

HOW TO SUBMIT:

- Label your paper nicely at the top with your name and topic title
- Submit your paper via the drop box before Sunday 8pm
- On Monday morning, we will merge the papers together into a single MS Word document and convert into a pdf
- This pdf doc will be uploaded on the class site so that students can print copies before the class or view the pdf electronically during class
- An electronic update will be sent once the upload has been done
- Please make sure to include the list of reference articles and/or Internet sources. It's ok if the reference list runs into the second page.

The time to present the paper to the class is 3 minutes, followed by a couple of minutes for questions. Papers from previous semesters are available for your review

MACHINE LEARNING APPLICATION IN FRAUD DETECTION.

INSTRUCTOR: DR. ZOYA KINSTLER, TA: TAKAYUKI IIDA

One in three organizations experienced economic crime (PwC, 2017). More so, instances and variability of such are increasing with continuing growth and sophistication in digital technology (Reuters, 2015). A typical organization loses at least 5% of revenues each year to fraud (ACFE, 2013), making detecting and mitigating fraud important. These issues are prompting companies to devise ways to cut losses, reduce the cost to organizational reputation and mitigate the risks to acceptable level to help meet business goals. One of such cases is PayPal using machine learning (ML) for fraud detection.

PayPal currently uses a combination of several ML algorithms: linear, deep learning and neural network. With these algorithms, machines combatting fraud are trained with a gargantuan amount of past data with decision boundaries to draw insights, classify and allow predictions on newly related occurrences. Fei-Fei Li, the head of Stanford's Artificial Intelligence Lab exemplified this when she began labeling the millions of raw images that a child might encounter by age three, and feed them to computers. So that by being shown vast numbers of labeled datasets with instances of say, a dog, the machine will learn from it like a toddler would shape its own rules of whether a particular set of digital pixels was, indeed, a dog (Fei-Fei & Karpathy, 2015).

This training process is primarily made possible because of big data, automated algorithms that have deep-rooted classical statistical inference foundation, and some human oversight. Big data is relevant because the better and more prior knowledge a machine has about occurrences, the better the accuracy of inferences or decisions made on related new data. Usually, the data is structured by assigning labels (outcomes) to groups of related attributes (features) of interest so that the machines can associate same labels with similar characteristics of new information. Statistical automation comes to play because models are built using high-level programming languages with samples from required data, meeting some statistical conditions, to generalize outcomes to out of sample dataset.

The generalization accuracy depends on the stringency of decision boundaries or statistical line of best fit drawn to differentiate between possibilities that are used to train the machines. PayPal initially started with linear algorithms years ago, with a boundary distinguishing between outcomes. However, in reality, fraud's variability and adaptive nature, together with grey areas between occurrences called for a more granular approach to deep learning that uses neural networks. This granularity requires far more data points and multiple layered features compared to linear algorithms, providing for better precision. Neural networks try to mimic human brain thought process by connecting the dot between data points with abstractions on multiple levels. Multiple decision boundaries were set to cater for grey areas between outcomes, with some lines curved to accommodate nonlinear cases in the real world (Knorr, 2015). With that said, humans still have to oversee the process because machines may not understand context best fitted for business goals, and cannot get the data, correcting the perception of robots replacing humans.

This approach to combatting fraud is not a silver bullet. However, PayPal has achieved tremendous success with it in reducing their risks and improving processes (Crosman, 2016) after evolving the approach over the years. Aside from the automation, flexibility, and scalability it provides, it also frees up employees to focus on other areas of business processes

MACHINE LEARNING APPLICATION IN FRAUD DETECTION.

INSTRUCTOR: DR. ZOYA KINSTLER, TA: TAKAYUKI IIDA

that can provide added value, in addition to also providing a better customer experience in terms of maintaining trust. The benefits from this approach outweigh any disadvantages. After given much thought to its cost-benefit, companies can learn from this to help improve their processes in meeting business goals.

References

Reuters. (2015). Fraud rates on online transactions seen up during holidays: study

<http://www.reuters.com/article/us-retail-fraud/fraud-rates-on-online-transactions-seen-up-during-holidays-study-idUSKCN0T611T20151117>

PwC. (2017). 2016 Global Economic Crime Survey

<http://www.pwc.com/gx/en/services/advisory/forensics/economic-crime-survey.html>

ACFE. (2013). Report to the Nations on Occupational Fraud and Abuse

www.acfe.com/uploadedFiles/ACFE_Website/Content/rtnn/2012-report-to-nations.pdf

Fei-Fei, L., & Karpathy, A. (2015). Deep Visual-Semantic Alignments for Generating Image Descriptions

<https://cs.stanford.edu/people/karpathy/cvpr2015.pdf>

Knorr, E. (2015). How PayPal beats the bad guys with machines learning

<https://www.infoworld.com/article/2907877/machine-learning/how-paypal-reduces-fraud-with-machine-learning.html>

Crosman, P. (2016). How PayPal is taking a chance on AI to Fight Fraud

<https://www.americanbanker.com/news/how-paypal-is-taking-a-chance-on-ai-to-fight-fraud>