ANALYTICS FRONTIERS CONFERENCE 2018

BROUGHT TO YOU BY:

TERADATA.

AI & ML

Lessons Learnt & real-world challenges

Avishkar Misra PhD
Practice Director – Data Science & Analytics for Americas

Founded in 2010 industry thought leader focused on **Business Results**.

Founded **in Open Source --** Eye towards **analytic outcome.**

We are the **SI business of Teradata.**

Who Is Think Big?

1st Big Data provider 100% focused around open source.

Apache Hadoop and cloud ecosystem integration.

6000+ employees (2018)

Full spectrum consulting, data engineering, data science & BI support.

So far we have delivered **2000+** current projects at any time **$700M + in Revenue**.

Data Science & Analytics

Analytic tools

Languages

Data sources

Analytic engines

Teradata Aster

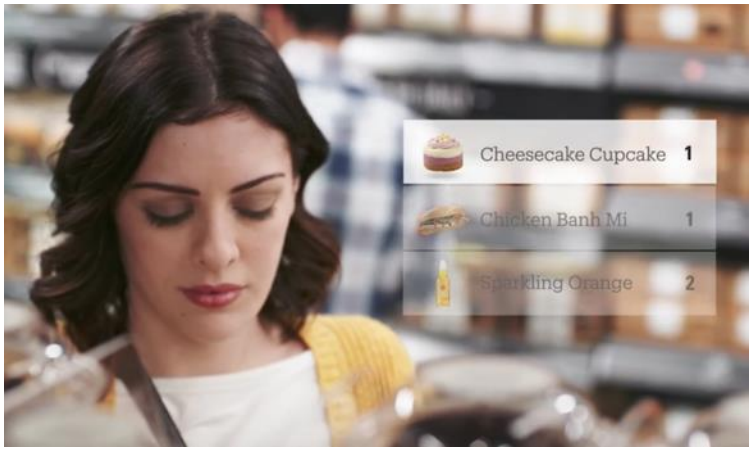

Five  lessons

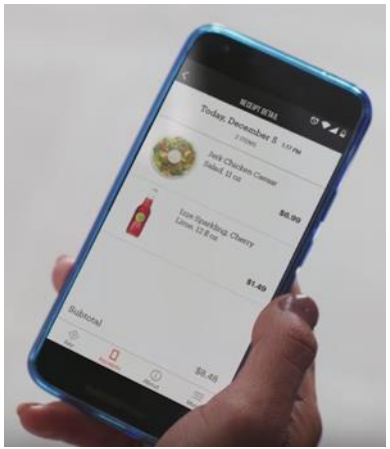
Amazon Go: No Checkout. Pick up and walk out.

source: amazon.com/go

TERADATA.

Amazon Go: No Checkout. Pick up and walk out.





source: amazon.com/go

Save time for customers

Improve customer service

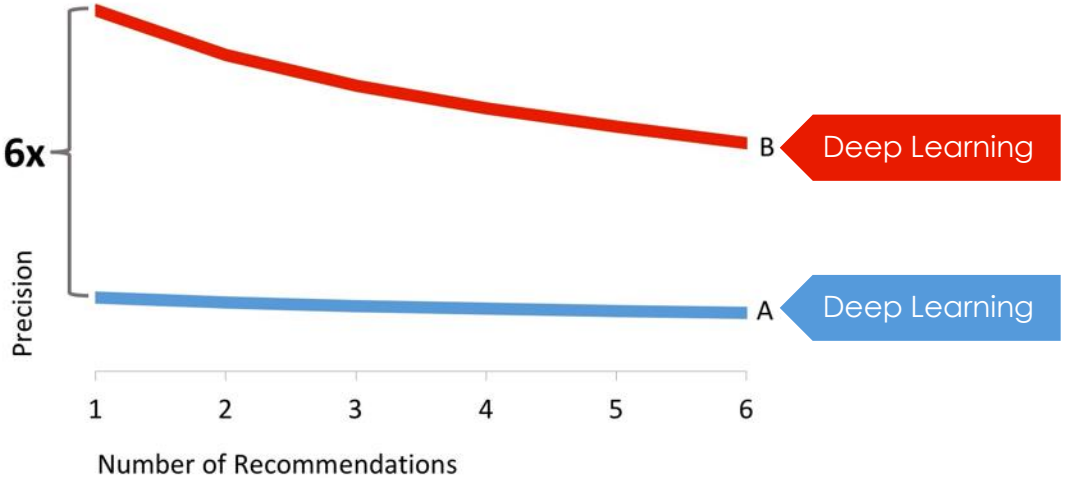
Prevent empty shelf events

Coupons 2.0

#1 – Flow-on benefits (& consequences) extend beyond the obvious.

TERADATA.

Deep Learning for Recommendations



A & B had identical neural network structure, size and training parameters.

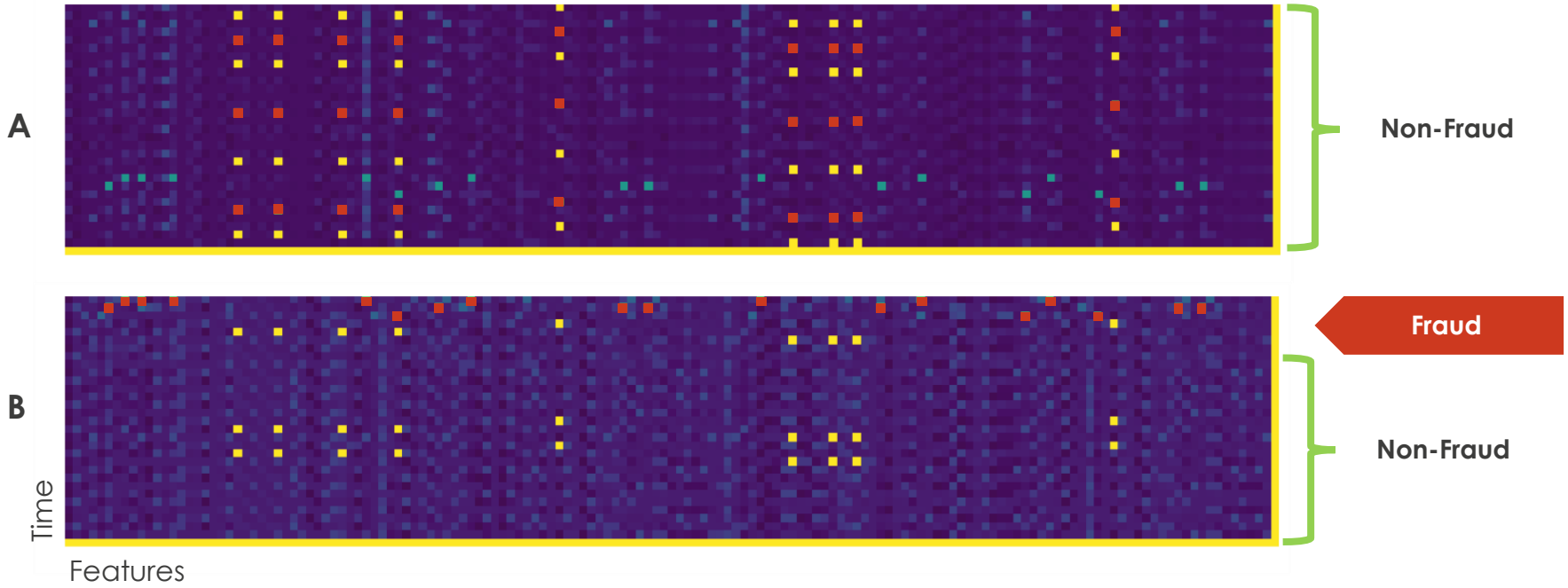
The difference is how the problem was formulated.

Details: Rybakov et al., *The Effectiveness of a Two-Layer Neural Network for Recommendations*, ICLR 2018
→ https://openreview.net/forum?id=B1lMMx1CW

#2 – Formulate and solve the right problem, even if it is not how its done now.

TERADATA.

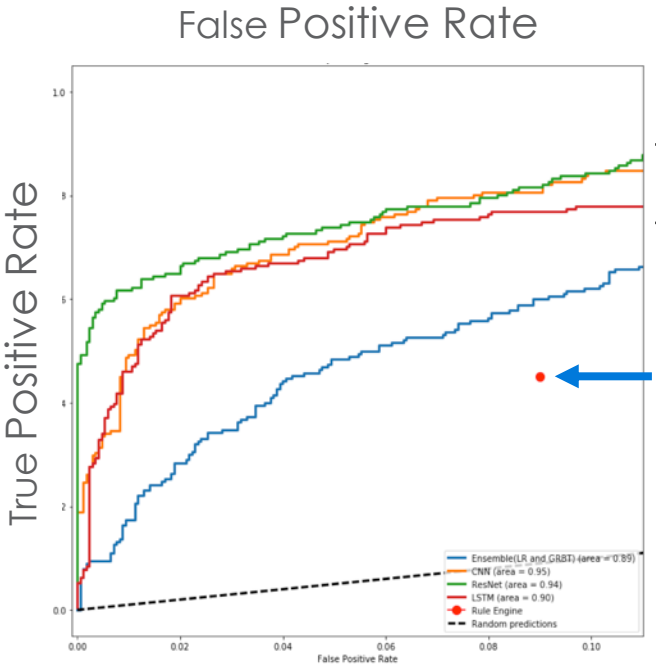
Deep Learning for Fraud Detection



Neural Networks (CovNets, LSTM, Autoencoders) to transaction images

TERADATA.

Deep Learning for Fraud Detection



Deep Learning

Machine Learning

Legacy Rule System

50% ↑ Fraud Detection Rate

60% ↓ False Positive Rates

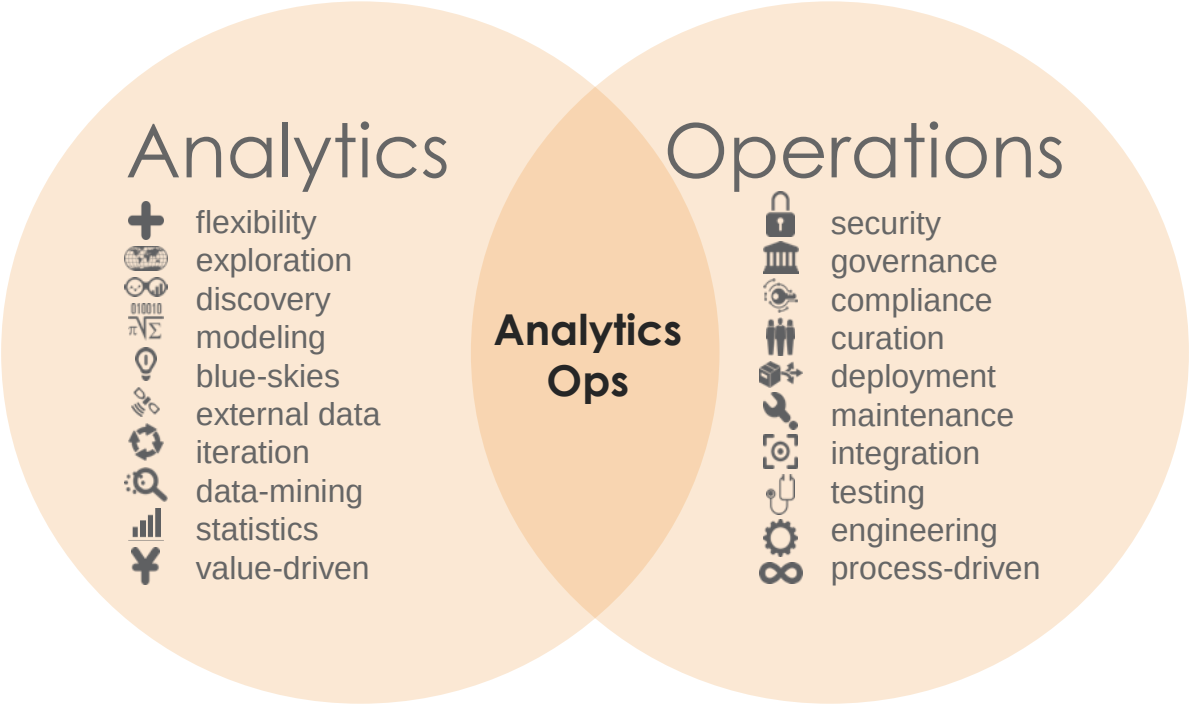
#3 – Avoid domain myopia and look for inspiration across domains.

TERADATA.

Across the board – Getting to Production

5 months
Average to develop, test, validate, deploy and scale new analytical models

2 weeks



#4 – Work backwards. Think production first. The lab is not enough.

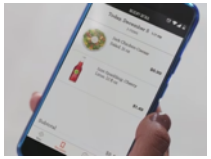
Across the board – Failure is inevitable

It ~~may~~ will **NOT** work the first time…

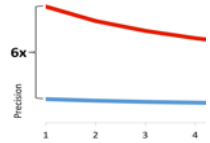
... or even the 4th time.

Mitigate the cost of failure.

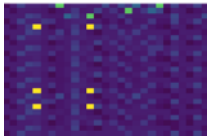
#5 – De-risk failure by speeding and scaling up experimentation



#1 – Flow-on benefits (& consequences) extend beyond the obvious.



#2 – Formulate and solve the right problem, even if it is not how its done now.



#3 – Avoid domain myopia and look for inspiration across domains.

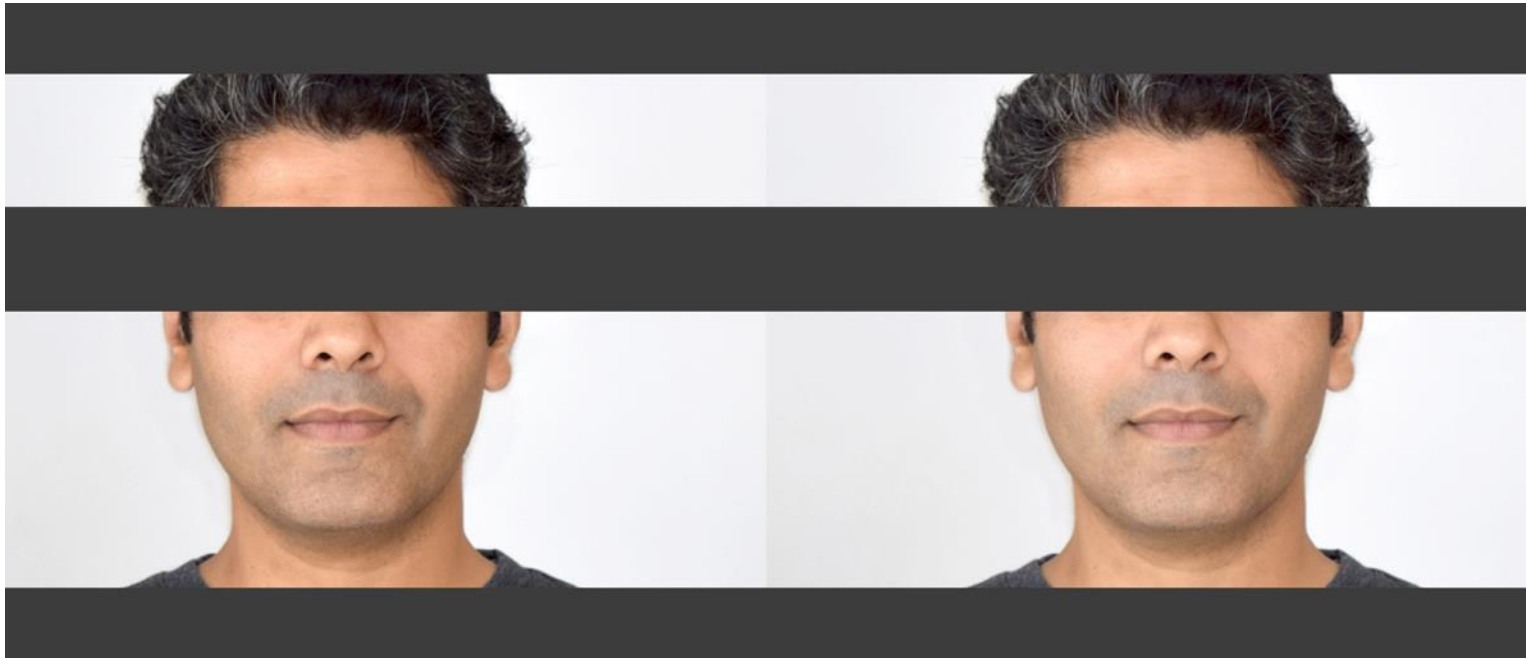
#4 – Work backwards. Think production first. The lab is not enough.

#5 – De-risk failure by speeding and scaling up experimentation

TERADATA.

Risks in adopting AI & ML

Biased Data and Poor Quality Control



One of these photos failed automatic quality checks

What can we do…?

#1 – Train Humans

More CS grads AND key decision makers in organizations.

#2 – Establish Standards

Processes and methods to verify and validate data & models holistically.

#3 – Automated Tools

Automated tools, verification and validation to surface potential data & model biases.

Lets talk…

Contact:

- Email: Avishkar.Misra@Teradata.com
- Blog: http://a-misra.com
- LinkedIn: https://www.linkedin.com/in/avishkarmisra
- Twitter: @AvishkarMisra

References:

Rybakov et al., *The Effectiveness of a Two-Layer Neural Network for Recommendations*, ICLR 2018, https://openreview.net/forum?id=B1lMMx1CW

Image Credits:

- Photo by felipe lopez on Unsplash
- five fingers by H Alberto Gongora from the Noun Project
- Danger by Gregor Cresnar from the Noun Project
- education by Rockicon from the Noun Project
- tools by Aleksandr Vector from the Noun Project
- Checklist by Ralf Schmitzer from the Noun Project
- amazon.com/go