UNIVERSITY *of* NICOSIA

Department of Digital Innovation

MSc in Blockchain and Digital Currency

BLOC 526 – Emerging topics in fintech

Session 9 – Machine learning techniques for fintechs

Dr. Periklis Thivaios CFA, FRM, BTRM

Learning objectives and expected learning outcomes

Summary

Session 9 introduces machine learning techniques in the context of fintech innovation. Even though it does not go into the technical depths of machine learning algorithms, this session aims to illustrate the opportunities for fintech and the ways with which the power of machine learning and big data can be harnessed for disruptive activity.

Learning objectives

- Conceptual understanding of machine learning, its roots and its differences from traditional analytics
- Supervised and unsupervised machine learning techniques for fintech applications
- Applications of machine learning to financial services and fintechs
- Caveats associated with the use of machine learning techniques

Expected learning outcomes

- Appreciate where and how machine learning techniques can be applied to financial services
- Develop an understanding of the benefits and downsides of machine learning techniques in the context of fintech

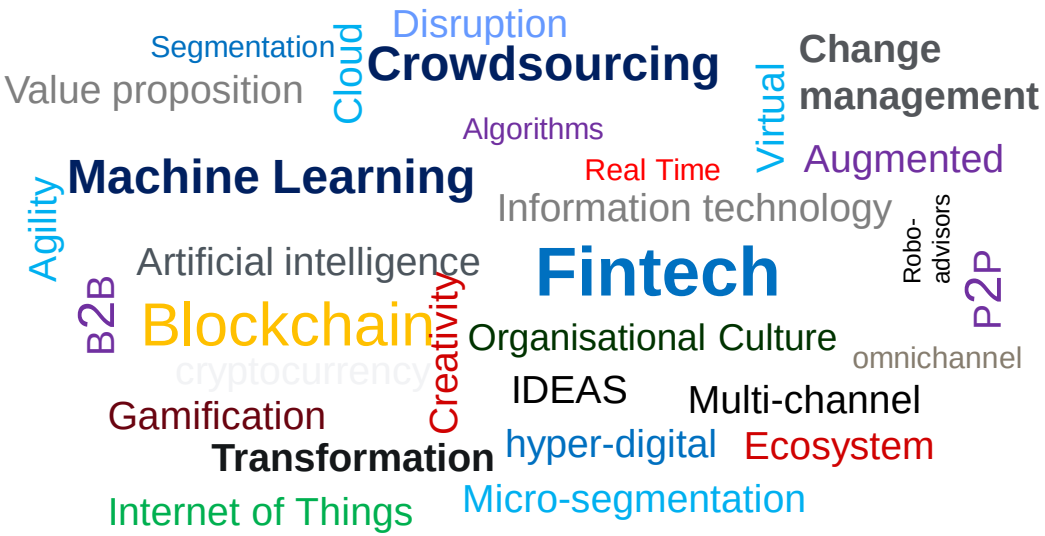
Agenda

- Introduction and recap
- Introduction to machine learning
- Machine learning and fintech
- Caveats pertaining to machine learning applications
- Concluding remarks

Agenda

- **Introduction and recap**
- Introduction to machine learning
- Machine learning and fintech
- Caveats pertaining to machine learning applications
- Concluding remarks

Today we will unpack one of the most common buzzwords in the fintech world



Agenda

- Introduction and recap
- **Introduction to machine learning**
- Machine learning and fintech
- Caveats pertaining to machine learning applications
- Concluding remarks

Question: What is machine learning?

Machine learning

Machine learning is an application of Artificial Intelligence (AI) typically used to produce insights and predictions from structured and unstructured data

Machine learning techniques are useful for:

- Making predictions: Identifying account holders that will leave the bank
- Look for interesting patterns: What products do my customers usually buy together?
- Search for groups in the data: What customer segments do I have?
- Highly complex processes: Modelling of non-linear relationships or modelling in high dimensions

Modelling

- Credit scoring
- Pricing
- Capital allocation
- Feature engineering
- Feature selection

Customer analysis (KYC)

- Segmentation
- Targeted marketing
- Sentiment analysis
- Consumption behaviour
- Willingness to pay

Model risk

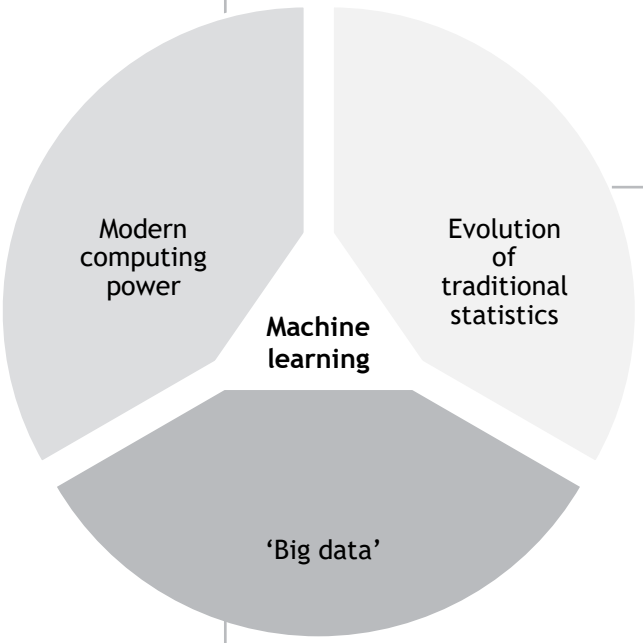
- Validation
- Back testing
- Monitoring

Fraud and financial crime

- Prospective fraud assessment
- Fraudulent pattern identification

Is machine learning that different from traditional techniques?

Yes and no!
Machine learning techniques build up on traditional techniques, advances in processing power and more data…



Modern computing power increase the speed and accuracy at which data can be processed and therefore has the ability to identify patterns humans will miss

Traditional Statistics

Machine learning

Supervised learning

- Regression
- Classification

Unsupervised learning

- Cluster analysis
- Dimensionality Reduction

- Deep learning
- Reinforcement learning
- Neural networks
- Generative adversarial networks

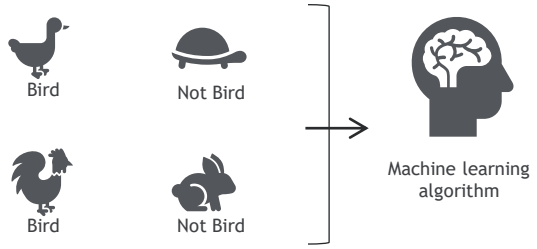
The growing quantity, velocity and variability of available data (structured and unstructured) enables machine learning

Supervised machine learning

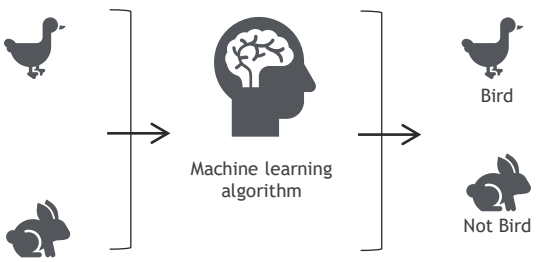
Supervised learning is the **machine learning** task of **learning** a function that maps an input to an output based on example input-output pairs. It infers a function from labeled training data consisting of a set of training examples

Supervised learning

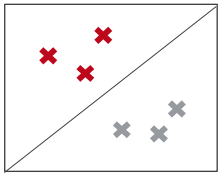
Step 1: Provide the machine learning algorithm categorised / labelled input and output data to learn from



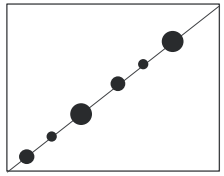
Step 2: Feed the algorithm new and unlabelled data to flag



Suited for the following problems



Classification



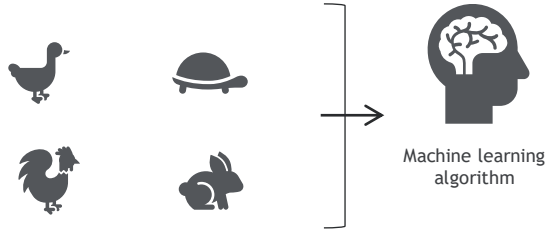
Regression

Unsupervised machine learning

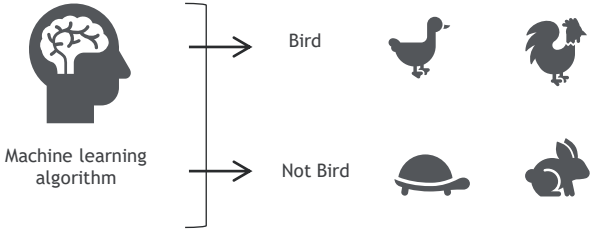
Unsupervised learning is a type of machine learning that looks for previously undetected patterns in a data set with no pre-existing labels and with a minimum of human supervision

Unsupervised learning

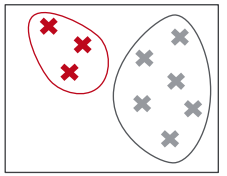
Step 1: Provide the machine learning algorithm uncategorised data to see what it finds



Step 2: Observe and learn from the patterns the machine identifies

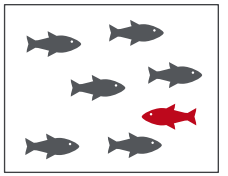


Suited for the following problems



Anomaly detection

Clustering

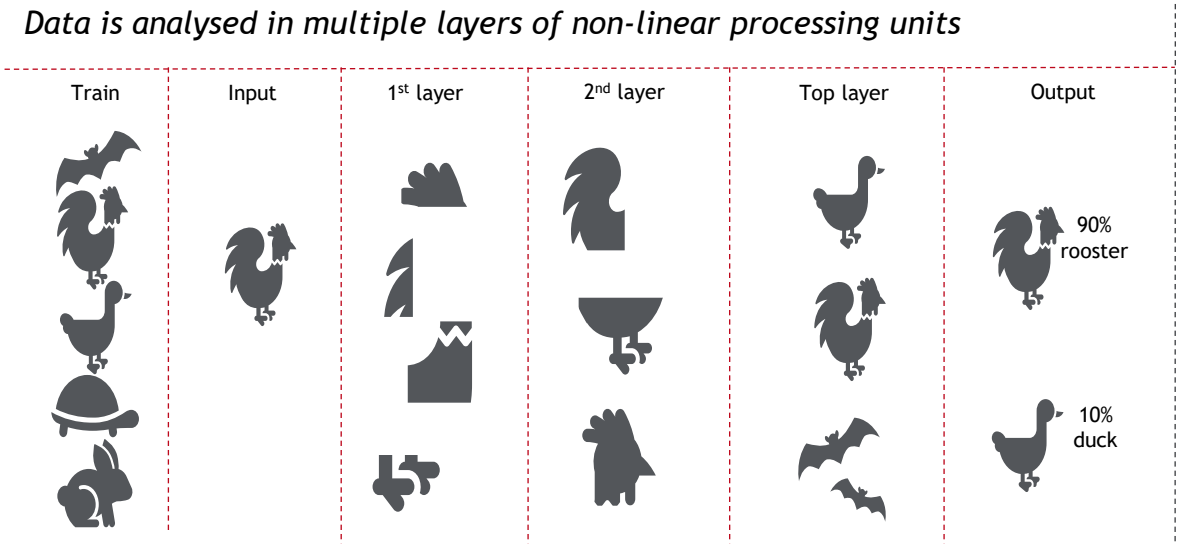


Deep learning

Deep learning is a function that mimics the workings of the human brain in processing data for use in detecting objects, recognising speech, translating languages, and making decisions

Deep learning

Data is analysed in multiple layers of non-linear processing units



Can be used in supervised and unsupervised problems

It is the intersection among research areas of neural networks, artificial intelligence, graphical modelling, optimisation, pattern recognition and signal processing

Reinforcement learning

Reinforcement learning is concerned with how software agents ought to take actions in an environment in order to maximise the notion of cumulative reward

Reinforcement learning

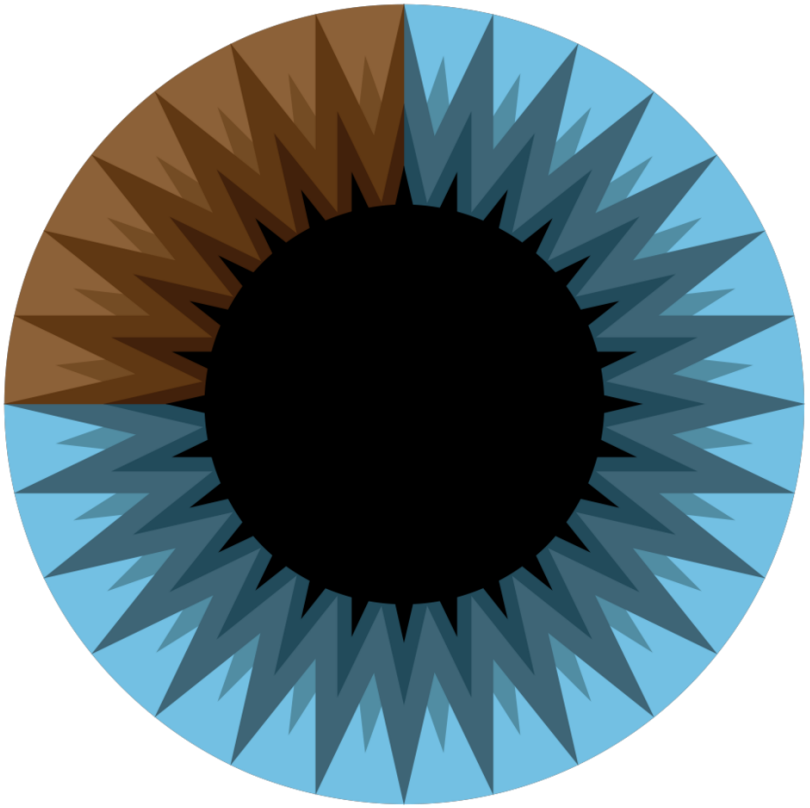
Allows the machine to learn a behaviour based on feedback from the environment. The behaviour can be learned once or can adapt as time goes by. It is used to learn a course of successive actions that maximise total cumulative reward / minimise the loss

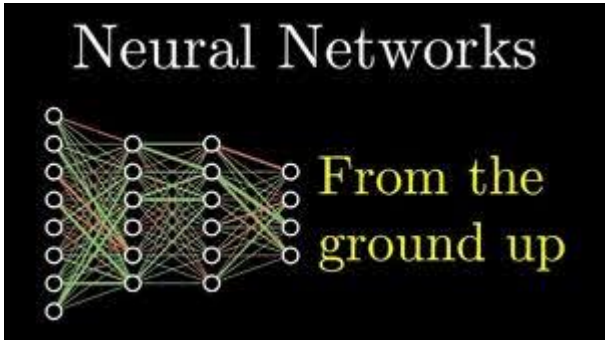
At each iteration, the algorithm is presented with a choice, based on the choice a reward is presented. The algorithm learns from the choice-reward combination in order to make better decisions in the next iteration

Often used in self-driving cars and process automation

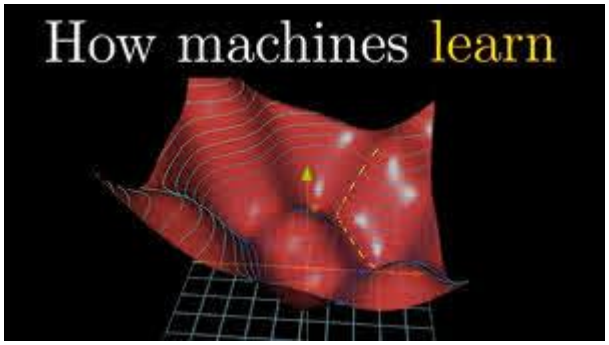
Want to learn more?

3 Blue 1 Brown youtube channel





https://www.youtube.com/watch?v=aircAruvnKk&list=PLZHQObOWTQDNU6R1_67000Dx_ZCJB-3pi



https://www.youtube.com/watch?v=IHZwWFHWa-w&list=PLZHQObOWTQDNU6R1_67000Dx_ZCJB-3pi&index=2

Takeaways

1. Machine learning is a 'new' toolkit for predictive analytics	▪ Enhanced methodologies, greater processing power and big data allow us to solve problems previously too cumbersome to do so

"What people call AI is no more than finding answers to questions we know to ask. Real AI is answering questions we haven't dreamed of yet"

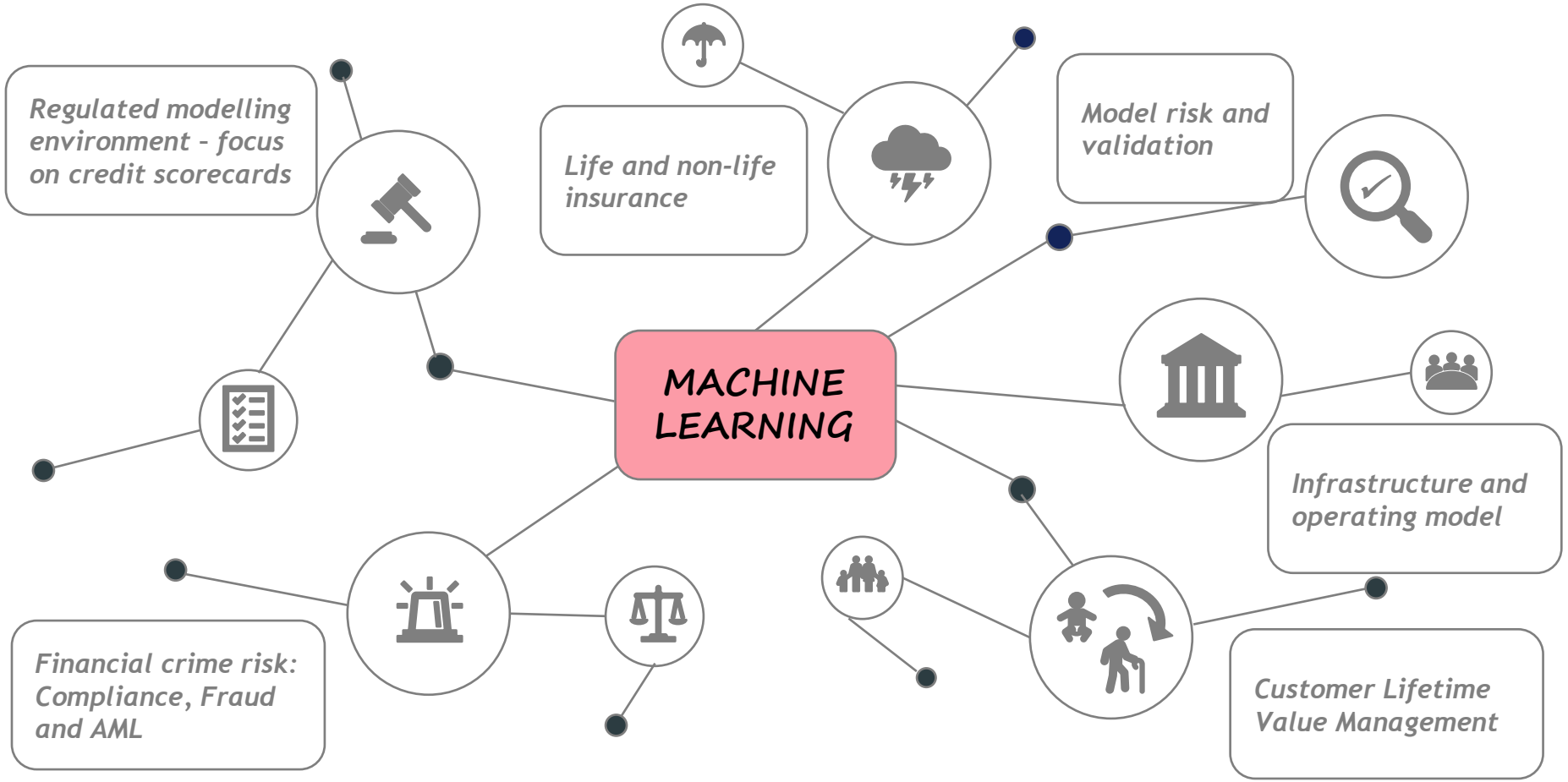
Tom Golway, Hewlett Packard

Agenda

- Introduction and recap
- Introduction to machine learning
- **Machine learning and fintech**
- Caveats pertaining to machine learning applications
- Concluding remarks

Question: Where do you think machine learning can be employed in financial services?

Advanced analytics can transform several things we do in financial services



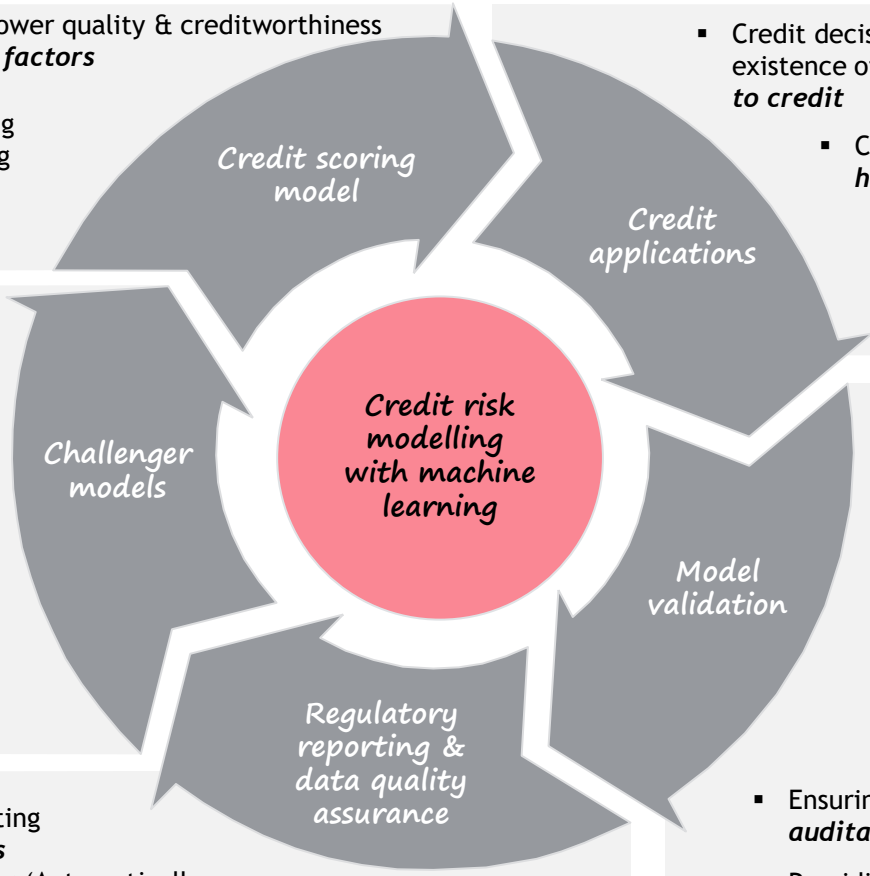
Machine learning algorithms combined with cheaper data storage and ever increasing computing power enables organisations to analyse data in ways previously not possible, paving the way for enhancing predictive analytics in the financial industry

Adoption area 1:
Regulatory credit risk modelling

- Captures more ***nuanced view*** of borrower quality & creditworthiness with assessment of more ***qualitative factors***
- Improved segmentation, feature engineering, variable selection, fitting techniques, regularisation (preventing overfitting), together with good transparency and less constraints on relationships with the target variable

- ***Tests*** against non-ML models with regulatory requirement constraints
- ***Selects variables*** and ***optimizes*** parameters in non-ML models using linear and/ or simple non-linear ML approaches
- Addresses risk associated with financial supervisors and the ***requirement of transparency*** for validation of risk models
- ***Improved accuracy*** of traditional scoring models (even when using traditional data)

- Incorporates demand of ***newer*** reporting requirements with ***greater resources***
- Efficient and effective data processing (Automatically ***identifies*** and ***amend*** anomalies in data (e.g. substantial errors, blank fields etc.)
- ***Lower cost*** and ***higher quality*** reporting

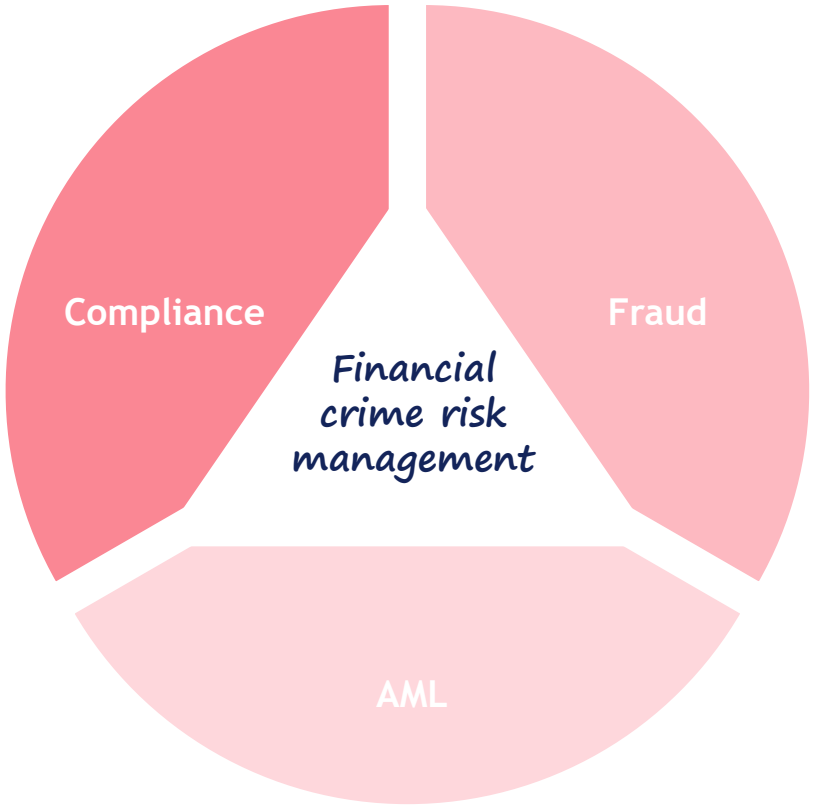


- Credit decisions with less dependence on the existence of credit history - greater ***access to credit***
  - Creditworthy borrowers with ***no credit history*** can be better measured
    - Cross-checking images in submitted identity documents - creditworthy borrowers with ***no credit history*** can be identified

- Polices outputs of primary models (internal and regulatory models) by using ***more***, ***real-time data*** to test model fit and generalisation
- ***Increased capacity*** to test scenarios and detect anomalous projections
- Management of ***model risks*** by ensuring that optimal models are developed

- Ensuring ***interpretability***, ***transparency*** and ***auditability*** relative to scope of application
- Providing powerful ***challenger*** tools - used when relaxing regulatory constrictions to provide ***benchmarks*** for constraint models

Adoption area 2:
Financial crime risk management



*ML has the potential to enable a **step change** in the **capabilities** used to **identify** and **manage** financial risks*

Process

Ability to process vast amounts of unstructured or poorly structured data

Identify

Ability to identify patterns and sequence to expose relationships which, on a standalone basis, do not provide pointers for concern

Flag

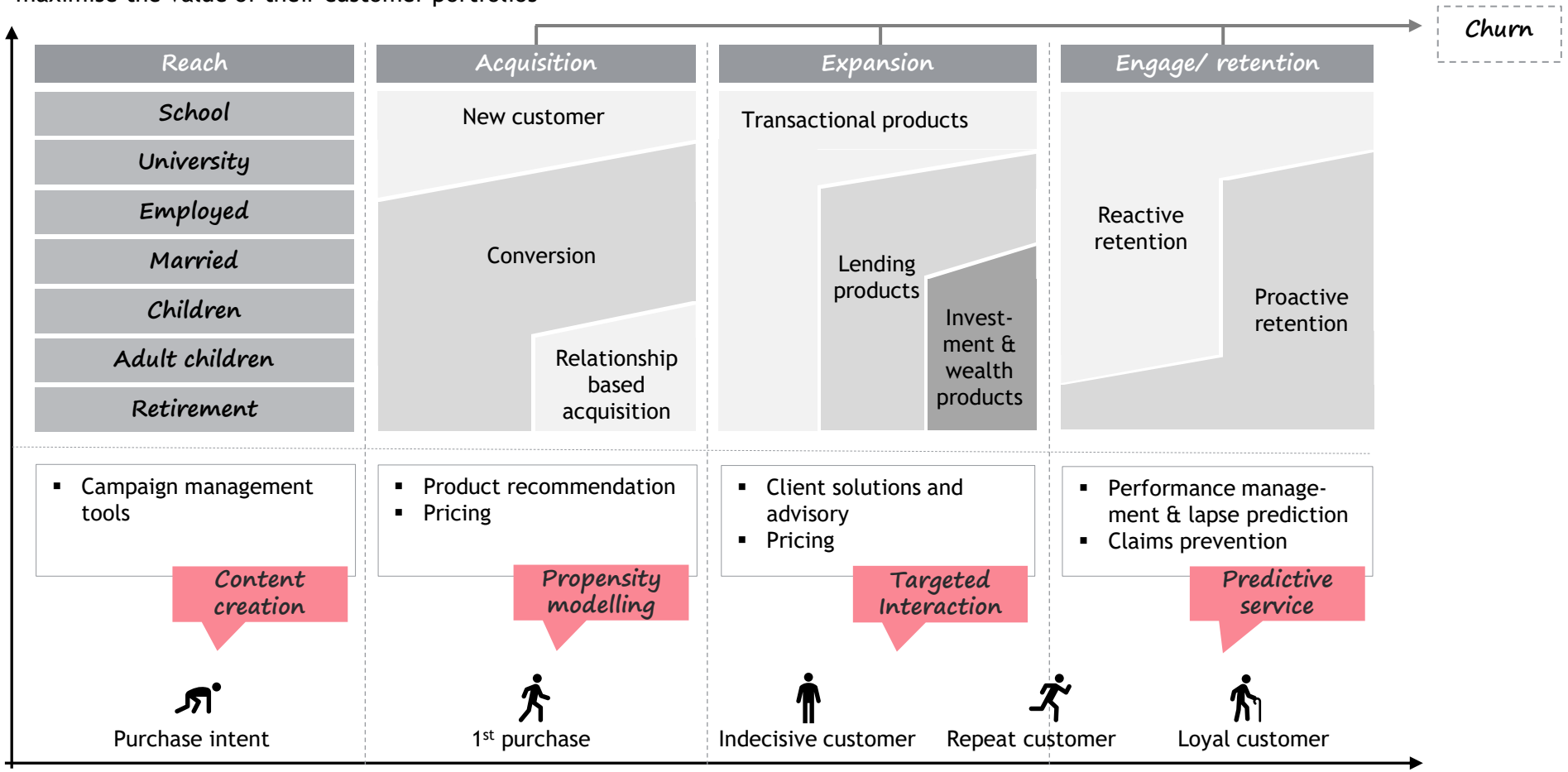
Ability to proactively flag suspicious accounts, account holders or transactions before fraud materialises

Monitor

Monitoring and reporting capabilities that can assist risk management and nurture the appropriate risk management culture

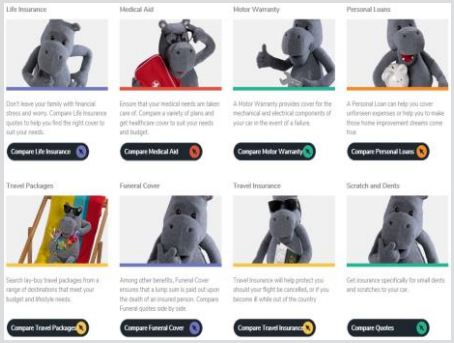
Adoption area 3:
Customer lifetime value management

Organisations can no longer sustain growth through acquisition alone; instead, they must manage the entire customer life cycle to maximise the value of their customer portfolios



Adoption area 4:
Life and non-life insurance

The modern client has available ***multiple insurance options*** at their ***fingertips***. The focus has shifted from interactions with local brokers / agents that sell fixed packages to ***online*** interactions ***tailored*** to client specific circumstances and requirements.



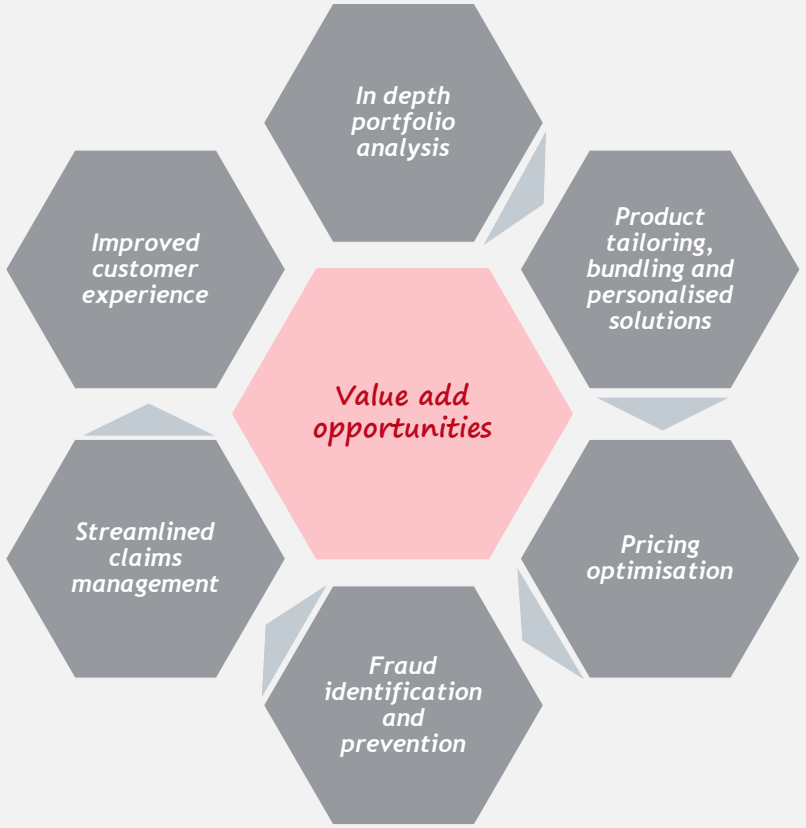
www.hippo.co.za

The modern client requires ***flexibility***, ***personalisation*** and ***speedy interactions*** and insurance organisations that are adapting to these requirements will gain the ***loyalty*** and ***business*** of these clients

As an industry that already gathers a large amount of data on it's customers there is large value-add for combining existing data with ***new and additional sources*** of information

Structured data	Type of loss, amount of loss, physician ID…
Spatial, graph	Accident location, work location, home location, common driving routes …
Text	Notes, diaries, medical bills, accident reports, depositions, social data, invoices…
Time series	Sequence of events, claim date, accident date, duration between events / action …

In order to keep up with industry developments, insurance companies are increasingly relying on machine learning to ***enrich their capabilities*** and ***improve their existing offerings***. Application areas of machine learning that provide the ***most value*** include:



Adoption area 5:
Validation of machine learning models and using machine learning for validation

Validations functions will be required to validate machine learning models and can further benefit from using machine learning models in the validation process

Model validation is a fundamental component of an organisation's model risk management framework, and a set of tools which are used to assess and manage model risk in support of the continued viability of the business, thus performing a crucial step in the model approval process

Model validation is the set of processes and activities intended to verify that models are performing as expected, in line with their design objectives and business uses. Effective validation helps ensure that models are reliable and sound through providing regular assurance that:

- Models are continually improved or evolved to remain relevant and fit for purpose
- Limitations are identified and appropriately addressed
- Models perform as expected and are adequately applied within scope, for the intended use as well as consistently throughout the organisation

Validation and use of machine learning models in a validation function

Validate
- As the adoption of machine learning becomes more common across different areas, validation functions will increasingly be required to validate machine learning models

Challenge
- Machine learning models serve as powerful validation tools against which modelling decisions and the performance of models can be tested, ensuring that the appropriate techniques are chosen at each step

Automate
- Automating certain aspects in the validation process will free up time for more detailed assessments of critical modelling decisions

Adoption of machine learning will require

Updates to validation process and framework
- It will be necessary for validation functions to adjust their validation process and framework, ensuring that additional considerations for the validation and use of ML models are included

Training staff and providing tools
- Validation functions have to ensure that their staff is sufficiently trained and have access to the necessary resources in order to efficiently use and validate machine learning models

Takeaways

2. Machine learning opens up opportunities for fintechs	▪ The fintech applications of machine learning are numerous ▪ Fintechs can become enablers for incumbent institutions, or disruptors where appropriate

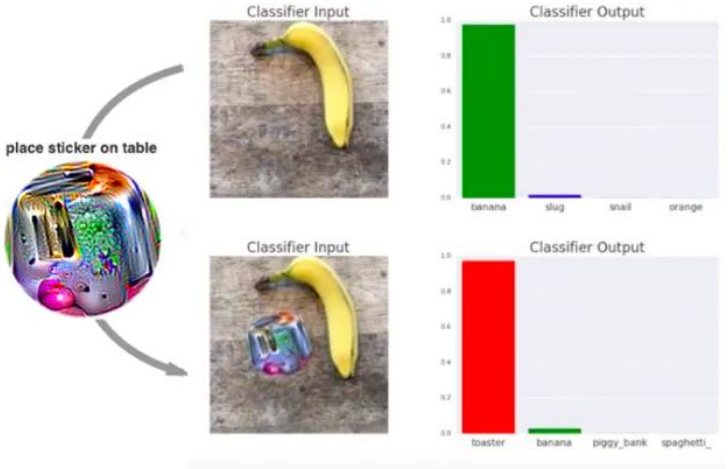
"Big data may be cheap to collect, but it's expensive if you don't know what you want to get out of it"

Ozge Yeloglu, Microsoft

Agenda

- Introduction and recap
- Introduction to machine learning
- Machine learning and fintech
- **Caveats pertaining to machine learning applications**
- Concluding remarks

Unsupervised machine learning can get things horribly wrong if left unsupervised…



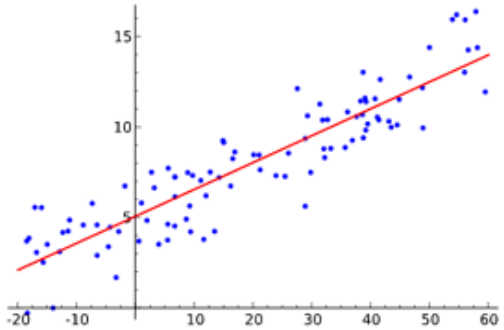
JOI ITO IDEAS 02.05.2019 08:00 AM WIRED

Supposedly 'Fair' Algorithms Can Perpetuate Discrimination

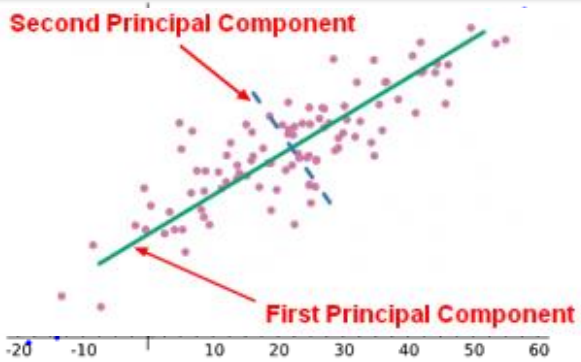
How the use of AI runs the risk of re-creating the insurance industry's inequities of the previous century.

Interpretability (or lack thereof) is probably the biggest challenge associated with the use of machine learning

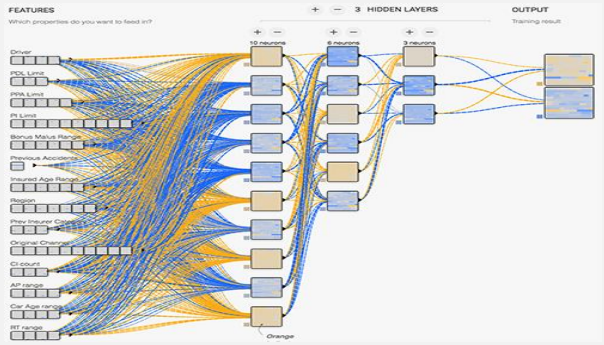
Simple regressions allow the logical validation of conclusions based on inputs



Principal component analysis is not widely used because of its lack of interpretability



Can we reject a loan or insurance policy on the basis of an incomprehensible model?



ML algorithms and techniques are powerful tools, but they must be harnessed in a responsible way to avoid detrimental effects…

Black box models	Using machine learning models without considerations of transparency, interpretability and auditability requirements will lead to models that ***don't work well in changing environments*** and can lead to the potential ***misuse of models*** when implemented, contributing to the ***lack of understanding*** and ***trust*** in models
Bias	Machine learning models developed on ***inappropriate*** data can reinforce or introduce unfair systematic biases The use of additional data sources are especially prone to introducing biases
Model Risk	Model risk introduced by machine learning models is often ***difficult to measure***, especially for more complex and less transparent algorithms
Social responsibility	The use of additional consumer data can introduce new risks for ***data privacy*** and ***security***. It is further difficult to predict how models using this data will affect the market in the long term, especially the ***impact of unintended systematic discrimination***

Takeaways

3. Machine learning should be applied with caution!

- Despite the possibilities, machines may not be as intelligent as we think
- Or, even worse, machines may be more honest than humans…

"Machine Learning is the process whereby if you don't like the results of a model, you shake it and get new results out of the same data"

Prof. David Taylor

Agenda

- Introduction and recap
- Introduction to machine learning
- Machine learning and fintech
- Caveats pertaining to machine learning applications
- **Concluding remarks**

Summary of takeaways

1. Machine learning is a 'new' toolkit for predictive analytics	▪ Enhanced methodologies, greater processing power and big data allow us to solve problems previously too cumbersome to do so
2. Machine learning opens up opportunities for fintechs	▪ The fintech applications of machine learning are numerous ▪ Fintechs can become enablers for incumbent institutions, or disruptors where appropriate
3. Machine learning should be applied with caution!	▪ Despite the possibilities, machines may not be as intelligent as we think ▪ Or, even worse, machines may be more honest than humans…

You don't need to know the techniques, the math or the software (well, ideally you do…)
But, if you want to succeed in the fintech world, you should at least use the term somewhere! (and ideally understand what it means…)

Any questions?

UNIVERSITY of NICOSIA

Questions?

Contact Us:

Twitter: **@mscdigital**
Course Support: digitalcurrency@unic.ac.cy
IT & Live Session Support: dl.it@unic.ac.cy