How Machine Learning improves Decision Making

Explanation of Machine Learning, when and where it will help you to considerably improve your decision making

Arshan Iqbal
Analytical Training Consultant, Decision Analytics

Gottfried Steiner
Business Consultant

Matthys Rossouw
Senior Business Consultant

experian

Overview

Introduction to Machine Learning

Use cases

When and when not to use it?

Experian approach to Machine Learning

The challenges our clients face with Machine Learning

Introduction to Machine Learning

What is Machine Learning?

An agreeable definition of Machine Learning is…

Machine Learning is the **delivery** of artificial intelligence so that computers are programmed to learn from **data.**

It tells the computer to look to identify for patterns, see how parts of the data correlates with other parts. Predicts likely outcomes to good level of accuracy, typically more predictive than traditional models

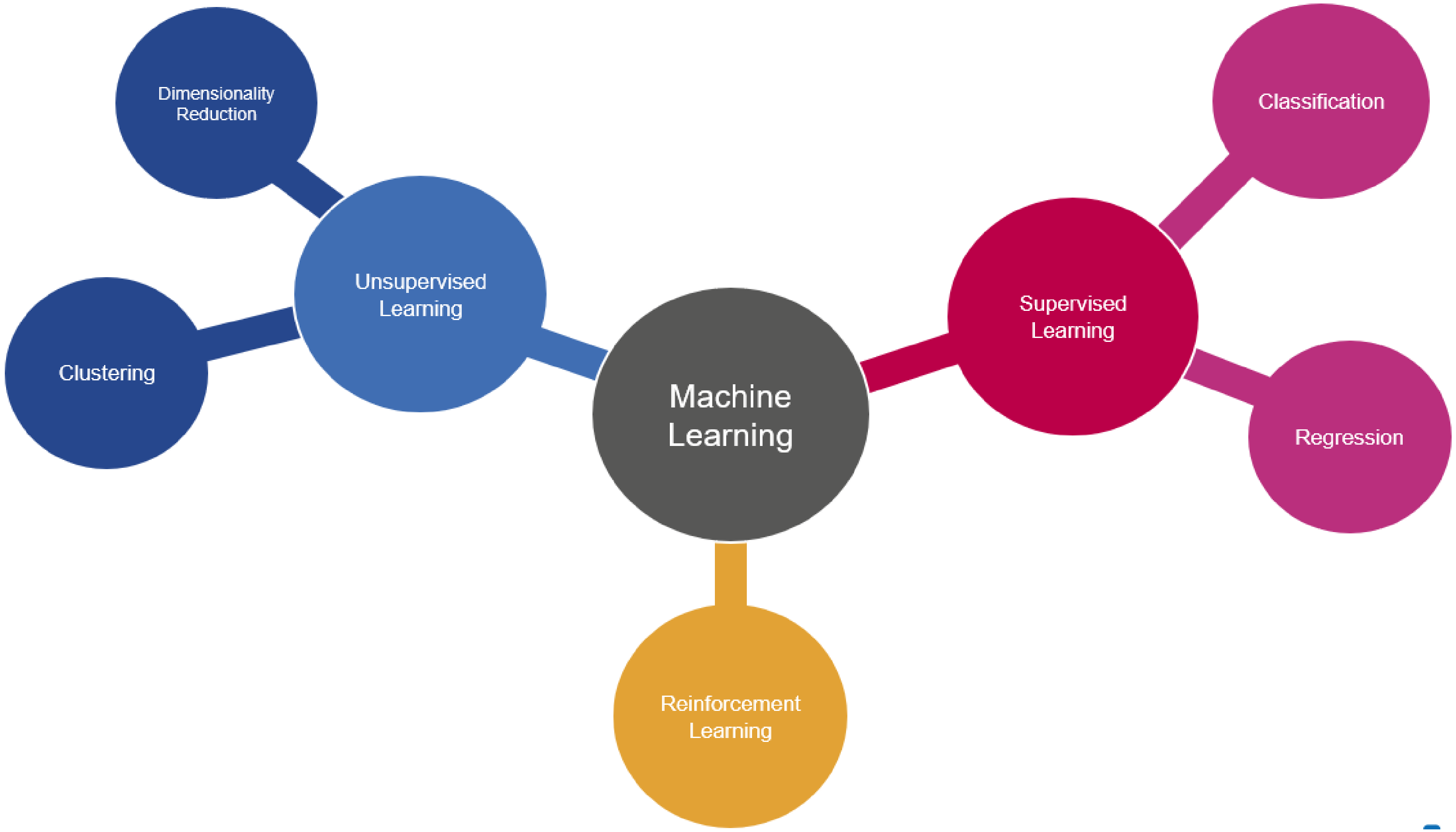
In relation to an agreeable definition of Artificial Intelligence…

Artificial Intelligence is the **theory** of programming computers to develop capabilities similar to human intelligence.

AI is the concept or philosophy, machine learning is doing it

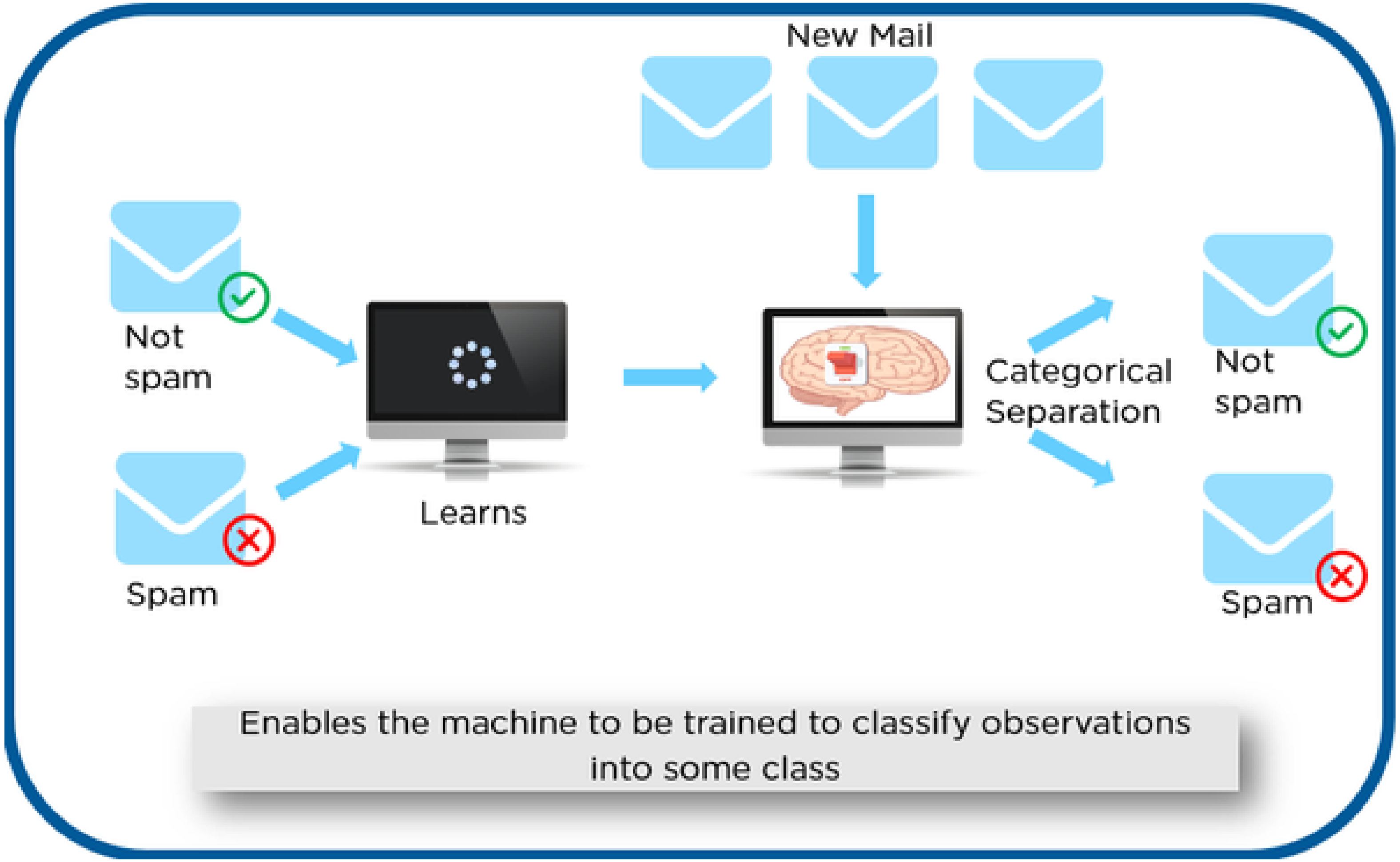
Introduction to Machine Learning

Types of Machine Learning



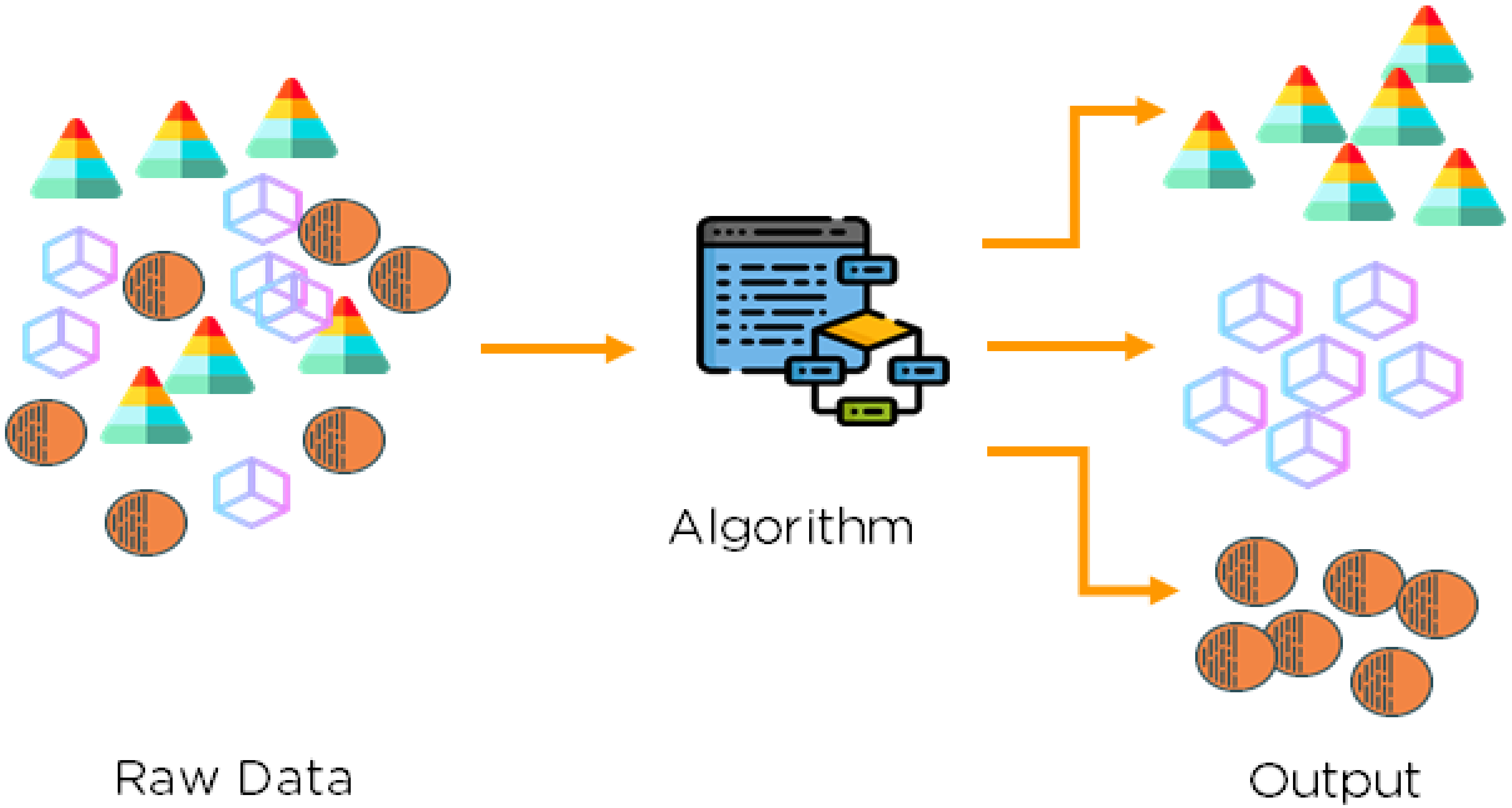
Introduction to Machine Learning

Supervised Learning



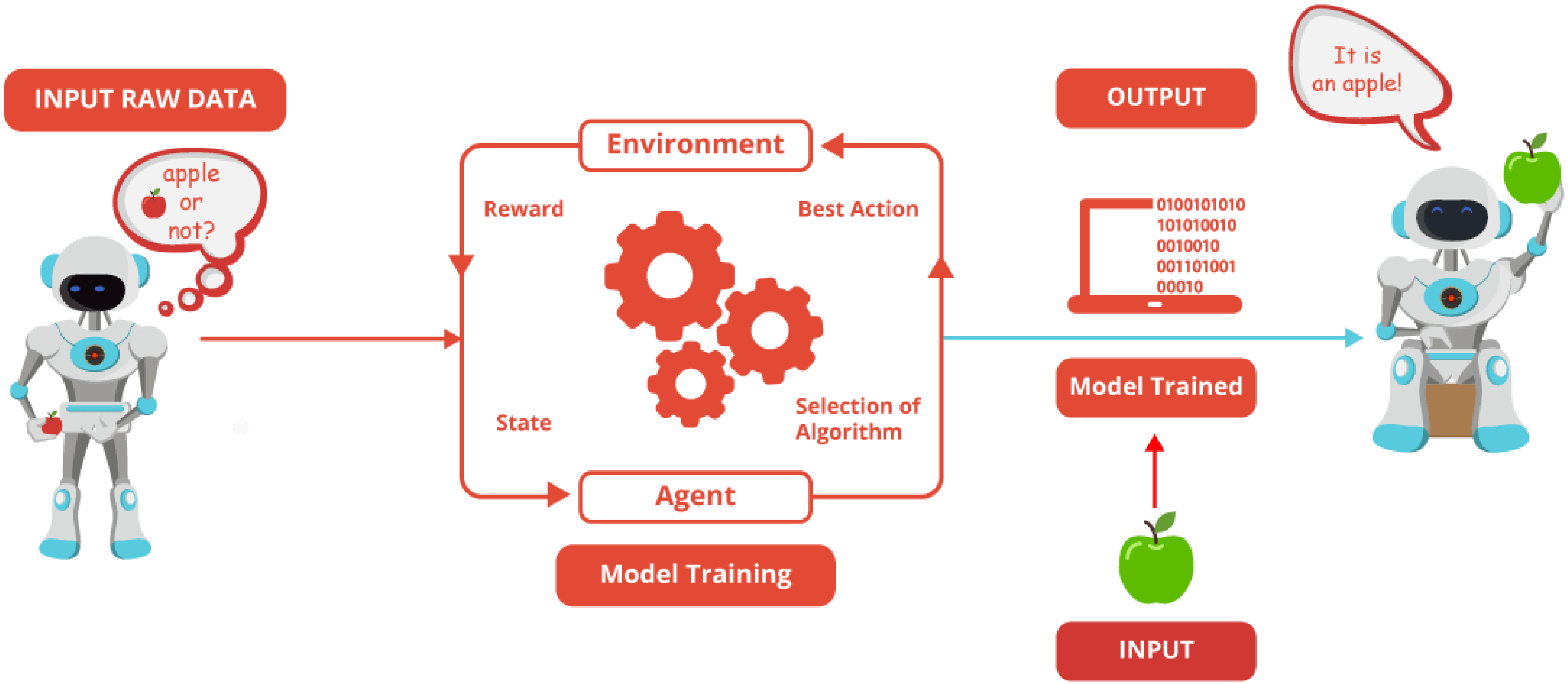
Introduction to Machine Learning

Unsupervised Learning



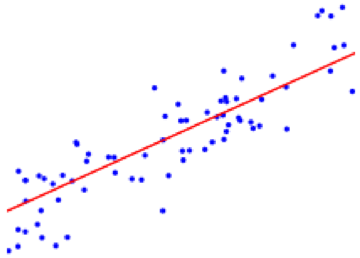
Introduction to Machine Learning

Reinforcement Learning

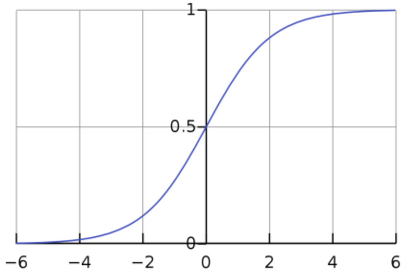


Introduction to Machine Learning

Linear Regression

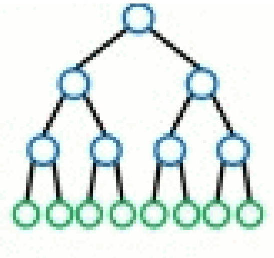


Logistic Regression

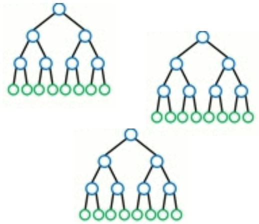


Traditional models

Decision Tree



Random Forest



Gradient Boosting

GBM

Advanced models

Introduction to Machine Learning

Why would you use advanced models?

- To achieve quicker model execution
- To unlock hidden data relationships and characteristics
- To achieve more accurate outputs
- To handle different data volumes, structures and formats
- To enable real-time and dynamic capabilities

Use cases

Indian banking and financial services company

An Indian banking and financial services company. It is the largest private sector lender by assets and the largest bank in India by market capitalization. They have deployed PCSM. The solution went live in December 2019.

Summary

Client wanted to implement their Python, R and H2O models quickly in PowerCurve. They developed the models using R Studio, Python and H2O Driverless AI Platforms.

Client stood up an instance of Microsoft Machine Learning Server (MSML) on their premise. They then deployed the R and Python models in PCSM and executed the models in MSML at run time.

For H2O models, the model jar files were exported from the H2O platforms and directly imported and executed within PowerCurve through the H2O ACE plugin.

The solution including the models was tested successfully in UAT and went live in Dec 2019.

Scorecards in Python, R and H2O are used for variable and scorecard creation.

ML use cases and benefits

Data preparation scripts in Python using Pandas library to convert raw to aggregated model ready variables

c. 250-300 Python, R and H2O models

Use cases:
Income estimation, credit score models and collection model

Algorithms:
Logistic regression, Decision trees, Random forest, Gradient boosting, Extreme Gradient boosting

Client Benefits:
Reduced model deployment time
Model recoding

Use cases

EMEA leader in Auto and Credit Financing

An EMEA leader in Auto and Credit Group financing, and an expert in debt collection which has gained wide recognition through its expertise, agility and sense of innovation serving over 300,000 customers.

Summary

PowerCurve® Strategy Management with integrated machine learning capabilities, enabling the company to deliver 1.5 million vehicle valuations every day, while ensuring further improvements to the accuracy of €1 billion asset re-sale valuations.

The business also wanted to ensure it could accurately and reliably offer consistent decisions on the 1.5 million quotations being requested every day, while at the same time, improving analysis and safe recovery of the €1 billion of assets being underwritten every year.

Client was focused on the improvement in analysis of vehicles' predicted cash surrender values, a major risk component of the medium-term commercial agreements amounting to around €1 billion in financial commitments for the company.

ML use cases and benefits

" Thanks to PowerCurve Strategy Management, Client Finance can call upon and deploy nearly 80 quotation models.

In addition to being able to quote future redemption values more precisely by using a large amount of available information, the challenge to be able to quote faster and in greater volumes to meet the expectations of new online partners. Today, the tool is handling about 1.5 million quotations every day."

— Director of Strategy and Data Science, Client organisation

Use cases

SAS Model operationalisation Case study

A leading bank in Latin America, part of a large multifunctional business group with businesses in banking, insurance, construction and mining among others, have operationalized SAS models in PCCM and PCSM

Summary

Client wanted to implement their SAS models quickly in PowerCurve. They had the original xmls of the models from SAS along with the design license for SAS.

Their model development-deployment cycle was to test the model in SAS, approve and then export as PMML.

The delivery team successfully proved to the client the ease of PMML model ingestion all the way up to execution within strategies. The solution including the models was tested successfully in UAT and went live in May 2019.

It has been used as default probability to approve or reject customers. A combination of the three models is used for credit limit assignation, offers and pricing.

ML use cases and benefits

PMML models ML predictions & classifications:

- Default probability in credit cards in a specific timeframe. E.g. @12 months
- Propensity to increase the use of the credit card above x%. Example 30%
- Clustering of characteristics such as Credit Line, Experience, Utilization, and other variables

Reduced model deployment time:
The ability to export SAS models as PMML saved time in flow construction &segmentations and in testing.

Model recoding:
Previously recoded the models in PowerCurve. With the new ACE Framework, Client only had to construct their ML model once in SAS, export, convert and deploy. No recoding.

Use cases

Why are clients taking up Machine Learning?

- They have more data collected, stored and processed
- They want to pick up hidden characteristics and patterns in existent data
- They believe they can achieve more accurate results
- Analytics team have performed a proof of concept of advanced models over traditional models
- Once they observe an uplift in risk handling, they want to deploy in their strategy

When and when not to use Machine Learning?

When to use Machine Learning

- If the application has not been fully understood and requires much data exploitation, advanced models are needed to pick up meticulous patterns and relationships in the dataset. This is need for Machine Learning. **Example is face recognition on smartphone cameras**
- When you have new trends or changes in a population, bank may want to update their credit scorecard models to **better represent the data.**
- They believe a scorecard model may be more error prone and **Machine Learning models mitigate error better**
- Cybersecurity is becoming a huge challenge because fraud is becoming more sophisticated. Therefore, Machine Learning is especially useful in preventing fraud and identity theft
- Help develop automated chatbots for automated and efficient scorecard calculation and customer service

When and when not to use Machine Learning?

When to use Machine Learning

- Create pricing models
- Help set credit limits
- Help upsell, cross sell or down sell
- Supervised Machine Learning is mainly used in decisioning. It is responsible for making predictions using historical data

When and when not to use Machine Learning?

When not to use Machine Learning

- If a type of data or application has been well understood consistently over a long period or because the relationship and patterns existing in the data is limited, then you may see little benefit of advanced models. **Example is having a credit scorecard model with a 50-year long dataset in the same population (characteristics have not really changed or do not change readily)**
- Advanced models may only supplement credit risk decisioning and may be used alongside traditional models, not replace them
- There may be issues with overfitted (over representative) models
- Any models is as good as the data used to calibrate it. If data is bad, the advanced model will not provide any more benefit than the traditional model

Experian approach to Machine Learning

ACE plug-in

Import the model

A
The model is directly imported into the studio

B
Studio is pointed to the analytics repository where the model is published

Map inputs and outputs

Inputs and Outputs of the model are referenced within the studio and mapped to the decisioning Data Universe

Test and Validate

The model execution component is then simulated and tested to ensure accuracy and consistency within the model creation environment

Add to Strategy

Bring the model into the flow to apply analytical insight within the strategy

Simulate

The Strategy can be simulated across the wider dataset to assess the performance of the strategy with the model as part of the flow

Deploy

Deploy the resultant strategy file for further testing and operationalisation within the Decision Agent Environment

Experian approach to Machine Learning

ACE plug-in

"Less than 15% of organizations believe they have fully operationalized the output of their advanced analytics" IDC

Where you build

Platforms for Machine Learning

Microsoft ML
SAS
Open Python
Open R
DataRobot
Dataiku
H2O
PMML
TensorFlow
Etc.

How you code

Programming Languages

python
sas
R
julia
Scala

What you build

Learning algorithms
(mathematical methodologies)

Decision Tree,
Random Forest,
Gradient Boosting,
Neural Networks,
Support Vector Machine,
Naïve Bayes,
Etc.

…What are clients doing with it now?

Teams/technologies/processes in very different places within the organisation: **insight-to-action cycle disjointed**

Hardcoding models into the operational decisions, requiring more skill and effort and imposes limitations

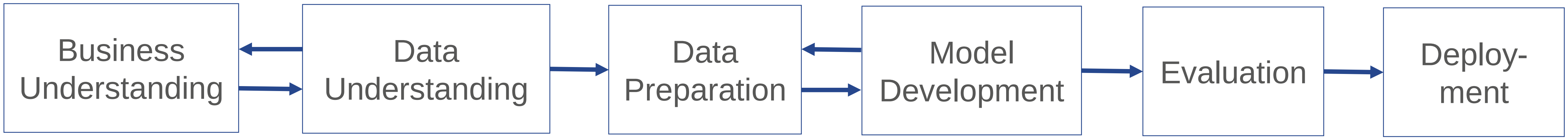
ML models not always explainable: in a regulated environment, **decisions become difficult to explain**

Changes to decisions made by machine learning sit with data scientists: **insight-to-action cycle becomes inflexible & slow**

experian

Experian approach to Machine Learning

Our approach to implementing Machine Learning models



Challenges clients face with Machine Learning

Challenges

- Explainability – Machine Learning is still considered as “black box”
- Putting Machine Learning models in production (deployment)
- Being able to monitor performance of Machine Learning models after deployment
- How do you process different and alternative data sources?
- Dependence with technology, computational power and storage

Challenges clients face with Machine Learning

Challenges

- Man power – many analytical resources within institution/bank. Customers can scale up/down based on resource needs. We manage those resources for clients
- Expectation of what machine learning can do – performance of machine learning is going to be different fraud against credit scoring
- Regulatory restrictions