Responsible AI

Delivering Data Science Safely at Scale

Tom Cronin

Head of Data Science

Tola Alade

Data Scientist

LLOYDS BANKING GROUP

1776

Sampson Lloyd II
1699 - 1779

Who are we?

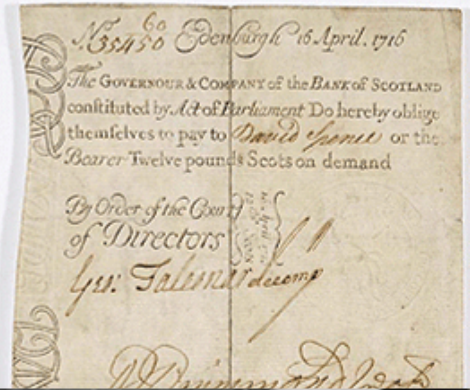
What is Lloyds Banking Group?

We own and manage some of the UK's best-known and trusted names in financial services. Everyday we serve over 25 million customers, and we're visible the length and breadth of the land.

Through our brands, we have supported the British people and their businesses for the past 300 years. And now we're building on this rich heritage as we aim to become the best bank for customers. By putting our customers first, keeping things simple and working to make a difference together.

This also means doing things the right way. Keeping our word. Earning our customers' trust. And playing a vital role in strengthening the UK economy.

Our purpose is simple: to help Britain's people, businesses and communities prosper.

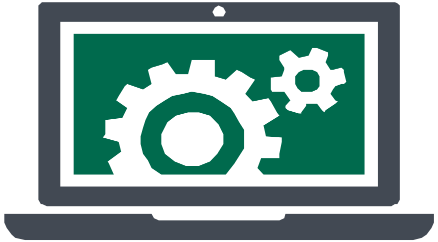

Edenburgh 16 April 1716
The Governour & Company of the Bank of Scotland
constituted by Act of Parliament Do hereby oblige
themselves to pay to David Spence or the
Bearer Twelve pounds Scots on demand
By Order of the Court
of Directors


Lloyds Cashpoint
How to operate Cashpoint

We are Group Transformation

Virtual Assistants

Data Science &
Machine Learning



Robotic Process Automation

We are the Applied Science Group

What is Responsible AI?

Data Ethics & Fair Models

Treating customers fairly through ethical use of data and models

Explainable AI

Enabling transparent Machine Learning models and tools to deliver better outcomes

Lifecycle & Governance

Embedded governance & control to deliver AI products safely and at scale

Explainable AI

XAI

AI is everywhere…

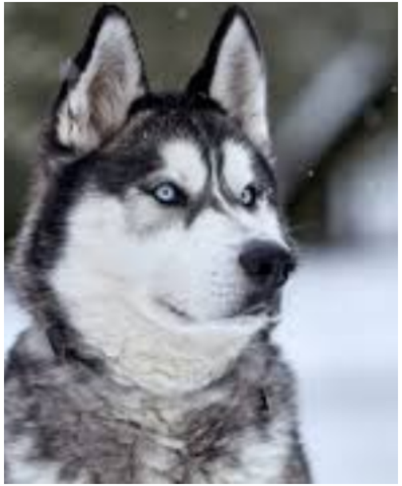
XAI is the ability to explain a machine learning model prediction

LLOYDS BANKING GROUP

Data → Black box Model → output → Approve loan application

XAI as a debugging tool

Predicted: wolf
Actual: wolf

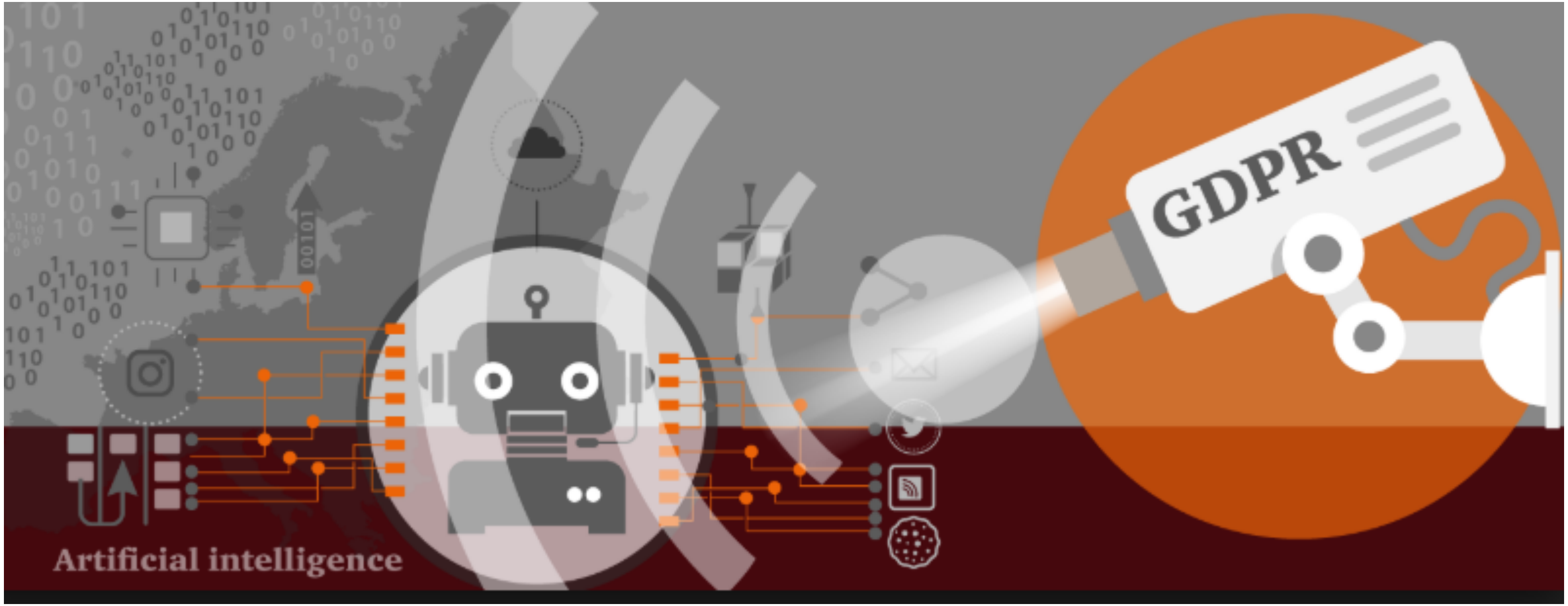


Predicted: wolf
Actual: husky

Predicted: husky
Actual: husky

Snow detector

Compliance

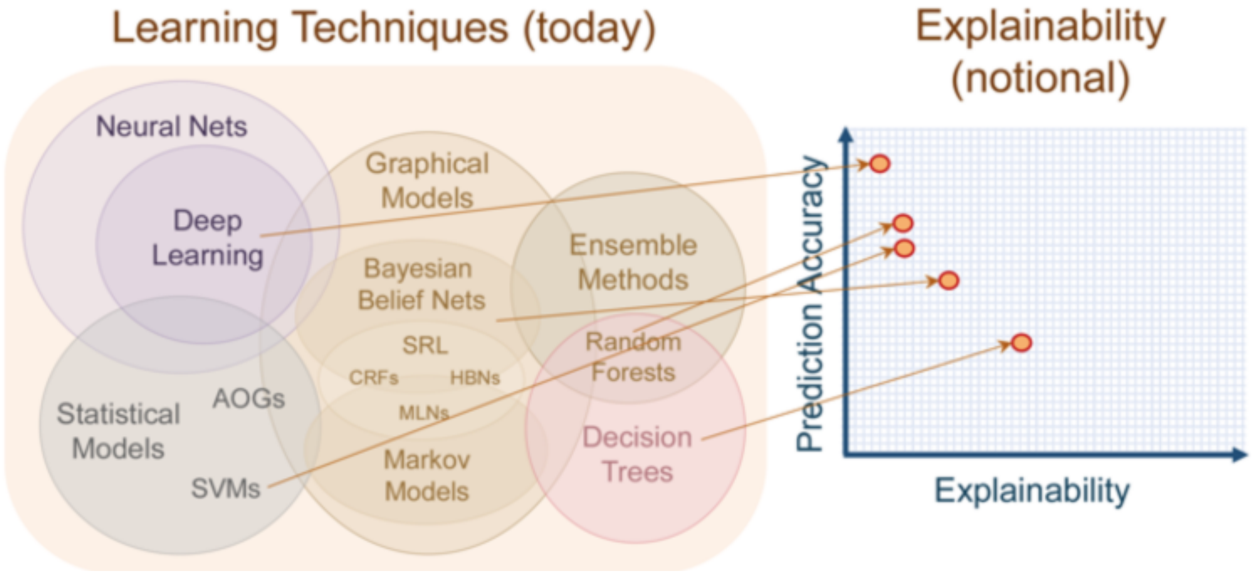


XAI aims to increase trust in models

Debit balance
CCJs
Number of missed payments
Credit balance
Total credit commitment

Data → **Black box Model** → output → Approve loan application

Interpretability SOMETIMES comes at a cost...



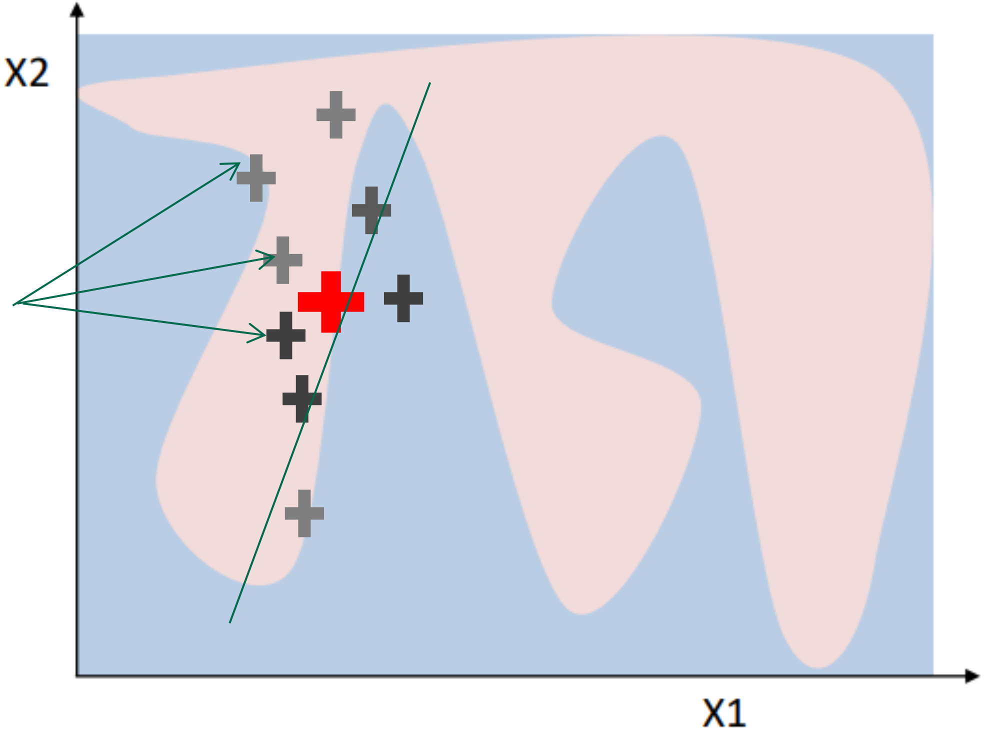
Source: http://nautil.us/issue/40/learning/is-artificial-intelligence-permanently-inscrutable

XAI Libraries

- LIME : https://github.com/marcotcr/lime
- Anchor: https://github.com/marcotcr/anchor
- SHAP : https://github.com/slundberg/shap
- ELI5 : https://github.com/TeamHG-Memex/eli5/tree/master/eli5
- Skater: https://github.com/datascienceinc/Skater
- sklearn-expertsys: https://github.com/tmadl/sklearn-expertsys

LIME Local Interpretable Model-Agnostic Explanations

Weighted explanatory samples



$$\emptyset_{\mathbf{1}} x_1 + \emptyset_{\mathbf{2}} x_2 + \ldots\ldots + \emptyset_{\boldsymbol{n}} x_n = \hat{\boldsymbol{y}}$$

$\emptyset_{\boldsymbol{i}}$ = coefficient of feature i
$\hat{\boldsymbol{y}}$ = estimate of $\boldsymbol{y}$
$\boldsymbol{y}$ = predicted outcome of the black box model

Anchor: Model Agnostic based on if then rules

	If	Predict
adult	No capital gain or loss, never married	$\leq$ 50K
	Country is US, married, work hours > 45	$>$ 50K
rcdv	No priors, no prison violations and crime not against property	Not rearrested
	Male, black, 1 to 5 priors, not married, and crime not against property	Re-arrested
lending	FICO score ≤ 649	Bad Loan
	$649 \leq$ FICO score ≤ 699 and $\$5,400 \leq$ loan amount $\leq \$10,000$	Good Loan

Source: https://github.com/marcotcr/anchor

SHAP Shapley Additive exPlanations

The Shapley value for a certain feature *i* (out of *n* total features), given a prediction *p* (this is the prediction by the complex model) is

$$\varphi_i = \sum_{S \subseteq F\{i\}} \frac{|S|!(|F|-|S|-1)!}{|F|!} \left[f_{S \cup \{i\}}\left(x_{S \cup \{i\}}\right) - f_s(x_s)\right]$$

Difference in predicted value with and without feature i added in to some subset of other features

|*S*| = length of set of feature groupings (minus the feature *i* we are interested in)
|*S*|*!* = number of permutations of set S
|*F*| - |*S*| - *1* = number of features to be added after feature I
(|*F*| - |*S*| - *1*)*!* = number of possible ways the features can be added
F = number of features

SHAP Shapley Additive exPlanations

Features: {Age, Height, Weight, Smoker}   **Predicted class for instance X:** Diabetic

Combinations	Age	Height	Weight
1	No	No	No
2	Yes	No	No
3	No	Yes	No
4	No	No	Yes
5	Yes	No	Yes
6	Yes	Yes	No
7	No	Yes	Yes
8	Yes	Yes	Yes

XAI demo with SHAP

Illustrative example

Task:

Classify household mode of transportation

Base model:

sklearn Random Forest classifier

Explanation model:

TreeSHAP

LLOYDS
BANKING GROUP