TRUSTWORTHY AI

FOR STRATEGIC INVESTMENT DECISION REGARDING CREDIT DEFAULT SWAPS (CDS) USING HPC

26/Sep/2023

Caesar Wu

Content

LUXEMBOURG
LET'S MAKE IT HAPPEN

- Strategic investment Decision for Credit Default Swaps (CDS) Using HPC
- AI/ML is considered to be a black box, How to Trust?
- Prediction Model: GBM/XGBM, Transformer
- Dataset and Sub-dataset: Technology and Telecommunication Sectors
- Explanatory Experiment Results
- Results Analysis
- Conclusion and Future Research Direction

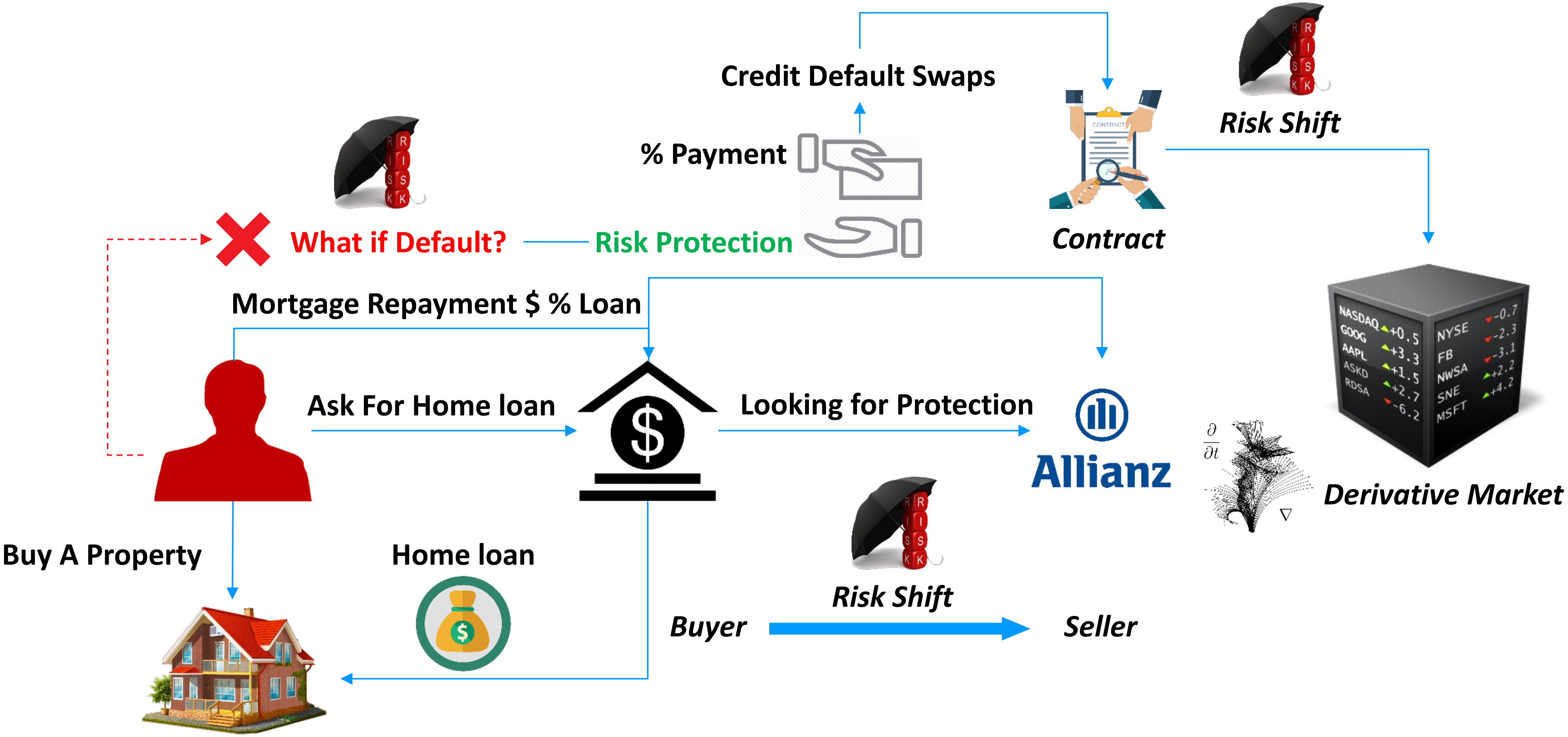
Research Problem

How Can We **Predict** CDS Price For Decision?

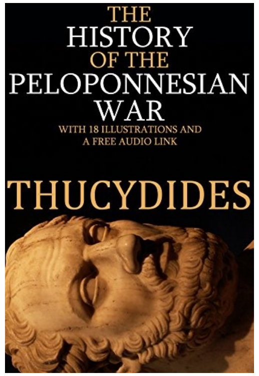
How Can We **Trust** the AI Results?

How Can We **Explain** the Result?

What is CDS?


Credit Default Swaps
Risk Shift
% Payment
Contract
What if Default?
Risk Protection
Mortgage Repayment $ % Loan
Ask For Home loan
Looking for Protection
Allianz
NASDAQ +0.5
GOOG +3.3
AAPL +1.5
ASKD +2.7
RDSA -6.2
NYSE -0.7
FB -2.3
NWSA -3.1
SNE +2.2
MSFT +4.2
Derivative Market
Buy A Property
Home loan
Risk Shift
Buyer
Seller

What is Trust and Trustworthy AI (TAI)?



Thucydides
460-400 BCE

Thucydides' conclusion:
"The vital difference between winner and loser = Leadership quality"

What is Leadership quality?

1.) Ability to process massive amount of information or data (AI)
2.) Quickly deciding what to decide (Trustworthy)
3.) Carry action with a resolution

"The philosophers have only interpreted the world, in various ways. The point, however, is to change it."

$$\textit{Trustworthy AI} = \frac{\textit{Trust}}{\textit{Distrust}} > 0$$

Why Does TAI Matter?

True Person's speech

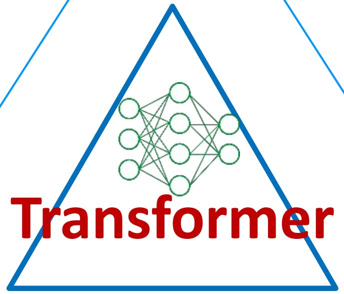
A few seconds of sample's voice

Imitation

Create

Not only Fake News, Image, and Voice

Voice



Transformer

Image

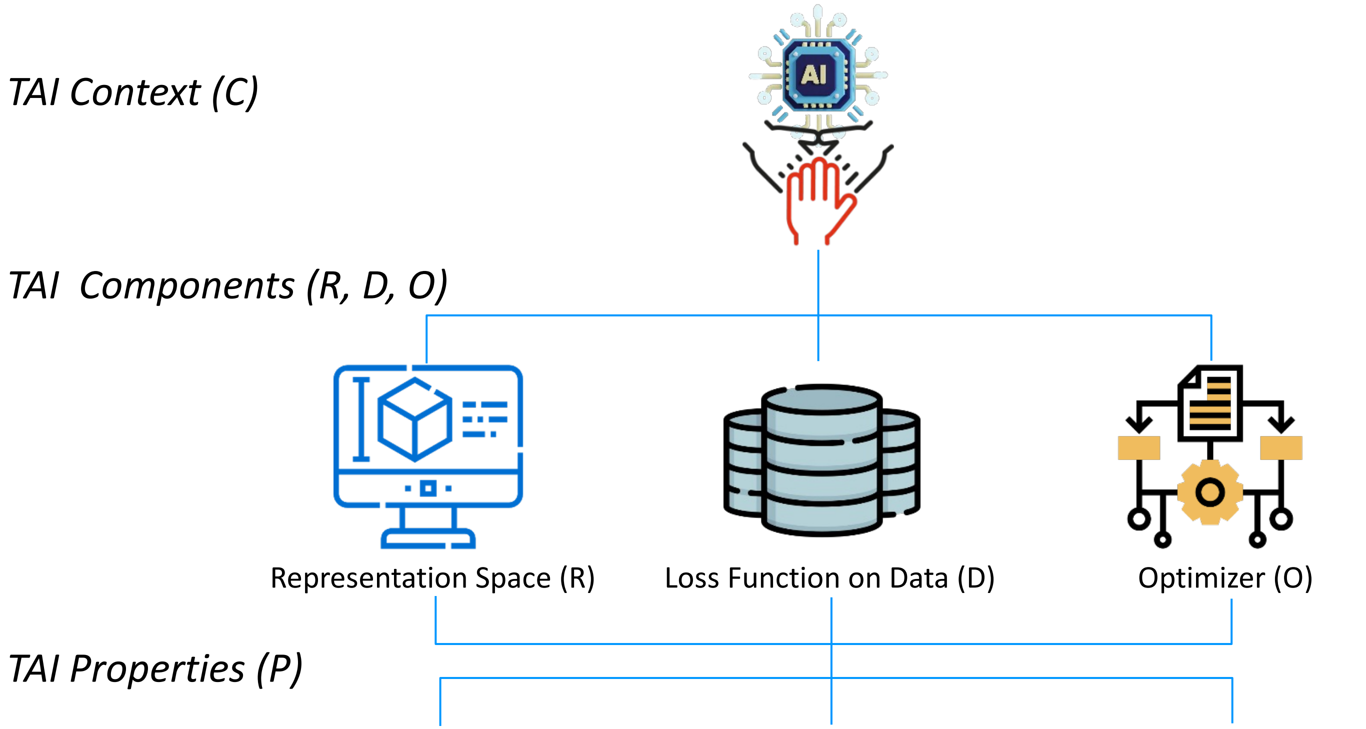
Text

Human cannot explain it

If choices are presented differently, people will make different decisions

Machine can manipulate human's mind

Create A Novel Framework For Trustworthy AI



Creation (Models)
- Explainability/Interpretability
- Transparency
- Fairness
- Justice

Evaluation (Data)
- Availability
- Usability (Data)
- Security/Privacy
- Accountability

Selection (Algorithms)
- Reliability
- Accuracy
- Robustness
- Reproducibility

How To Trust AI?

$$[\forall \boldsymbol{C}]\exists[\boldsymbol{R}, \boldsymbol{D}, \boldsymbol{O}] \vDash (\exists\ \boldsymbol{P})$$

or

$$\int T_{rust} = \sum_{i=1}^{k} p_i\,, \qquad p_i \in C$$

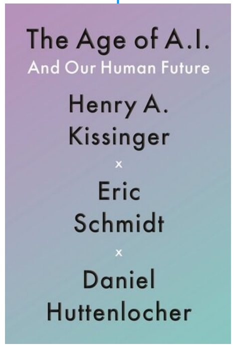
Why Some Trustworthy Properties?

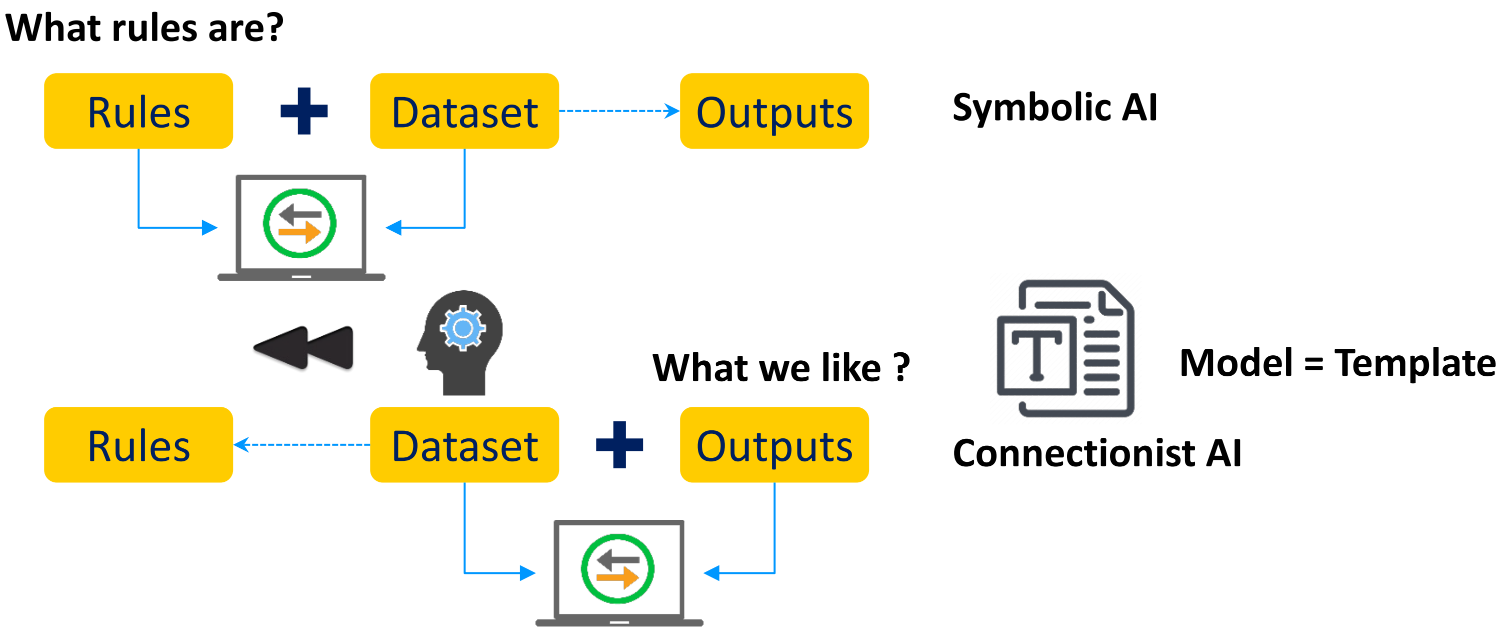
Explainable AI (XAI) for Decision Context

Decision Context or Applications:	*Examples:*
• Healthcare and medicine:	diagnostic assistance, treatment planning, drug discovery, patient remote monitoring (nearly all)
• Autonomous Vehicles:	self-driving, safety navigation, location mapping (security, reliability, robustness, accuracy)
• Finance and Banking:	fraud detection, risk assessment (security, reliability, accuracy)
• Customer Services and Support:	virtual assistant (reliability, robustness, usability)
• Cybersecurity:	threat detection, anomaly detection (security/privacy, reliability, robustness)
• Education	personalize learning (usability, availability)
• Legal and Compliance:	compliance monitoring, public order (justice, fairness, transparency)
• Criminal Justice and Public Safety:	court cases decision, legal aid (justice, fairness, transparency)
• **Environment and Conservation:**	wildlife protection, & monitoring, weather forecasting (reliability, accuracy)
• Social Media & Content moderation:	content filter (safety, accuracy, accountability)
• Government and Public Services:	resources allocation, public project bidding (transparency, accountability)
• Politics:	political campaign, election (transparency, fairness)
• HR:	candidate selection (fairness, justice)
• **Language Translation:**	language cross different culture (reliability, usability, availability)
• Manufactory	AI-Driven automation (reproducibility, reliability, robustness, usability, security)
• Emergency Response :	AI- Assist optimizing emergency response (reliability, availability, usability, accountability)

How to Frame TAI?

“The AI, then, did not reach conclusions by reasoning as humans ***reason****; it reached conclusions by applying the model it developed.”*

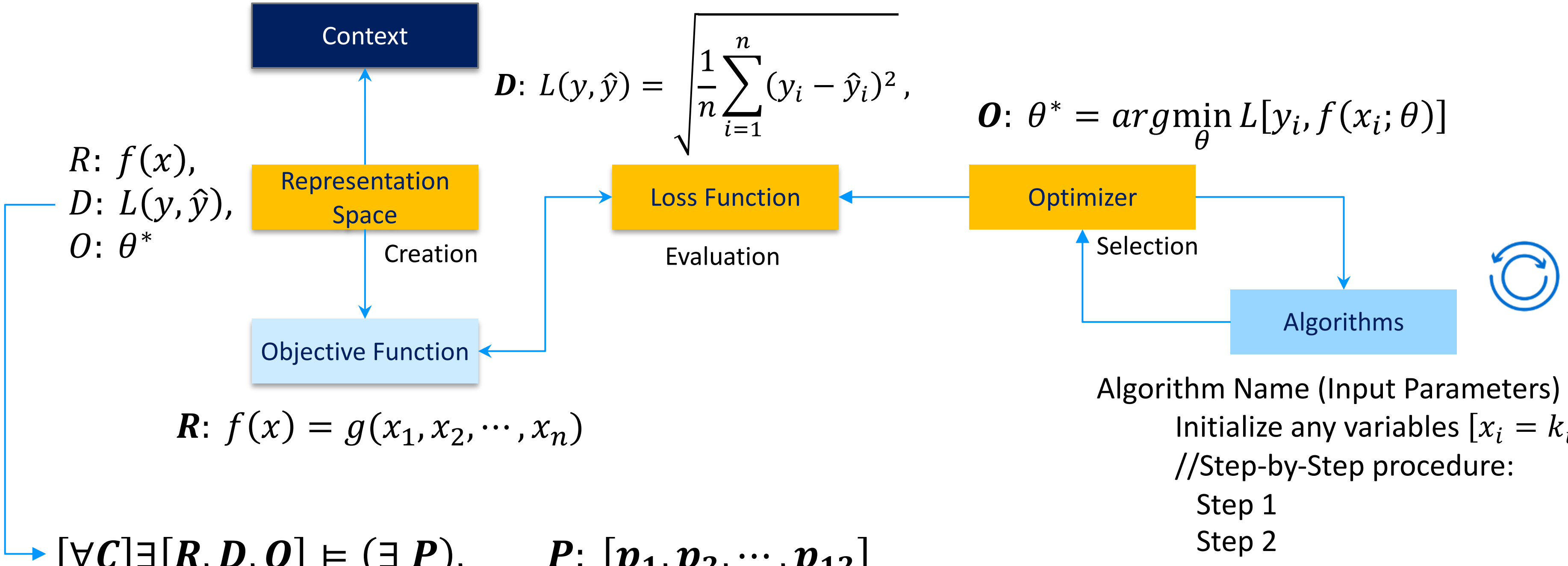




How To Implement TAI?

UNIVERSITÉ DU LUXEMBOURG

Context:
Decision Variables & Constrains
$\boldsymbol{C}: [x_1, x_2, \cdots, x_n] \in \mathbb{R}$



$\boldsymbol{D}: L(y, \hat{y}) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2},$

$\boldsymbol{O}: \theta^* = argmin_{\theta}\, L[y_i, f(x_i; \theta)]$

$R: f(x),$
$D: L(y, \hat{y}),$
$O: \theta^*$

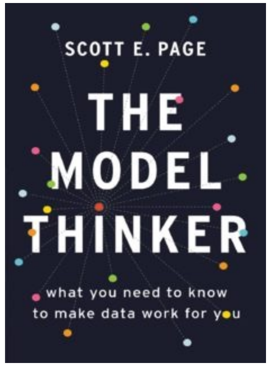
$\boldsymbol{R}: f(x) = g(x_1, x_2, \cdots, x_n)$

$[\forall \boldsymbol{C}]\exists[\boldsymbol{R}, \boldsymbol{D}, \boldsymbol{O}] \vDash (\exists\, \boldsymbol{P}),$ $\quad \boldsymbol{P}: [\boldsymbol{p_1}, \boldsymbol{p_2}, \cdots, \boldsymbol{p_{12}}]$

Algorithm Name (Input Parameters)
Initialize any variables $[x_i = k_i]$
//Step-by-Step procedure:
Step 1
Step 2
⋮

Types of Representation Space?

Seven Types or Three Groups of models



1. Descriptive

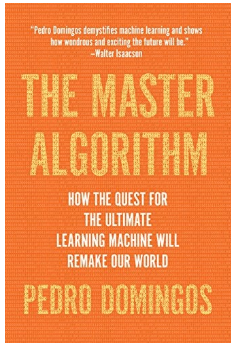
- Explain: To provide explanations for empirical phenomena (Create Value)
- Communicate: To relate knowledge and understand (Create Value)

2. Inference

- Reason: To identify conditions and deduce logical implications (Evaluation)
- Explore: To investigate possibilities and hypotheticals (Evaluation)

3. Predictive

- Design: To choose features of institutions, polices and rules (Selection)
- **Act:** To guide policy choices and strategic actions (Selection)
- **Predict:** To make numerical and categorical predictions of future (Selection)



Optimizer: Domingo's Five Schools of Thought on Machine Learning

Central Problem	Key Algorithms
Reasoning with symbols	**Decision Trees (If-then) or Tree-Based Model**
Analysing perceptual information	**Neural Networks (Perceptron, Deep Networks, Transformers)**
Managing uncertainty	Bayesian Networks (It has the statistics dependence on data)
Discovering new Structure	Genetic Programs (natural selection)
Exploiting Similarities	Nearest Neighbours (previous cases)

We Tested Two Different Models

1. Tree-Based Models

- Random Forest
- Gradient Boosting Machine (GBM)
- Extreme GBM (Xgbm)

2. Transformer Models

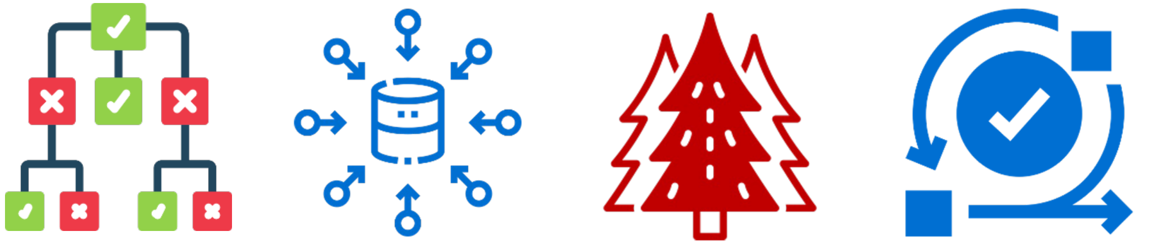
- Vanilla or Baseline model
- TimesNet,
- PatchTST,
- Crossformer

$$\nabla f = \frac{1}{n} \sum_{i} \nabla L(x_i)$$

Why **Tree-Based** Models?

- Won many Kaggle's Machine Learning competitions
- Tree-based models have a relatively long history
- Relatively easy to explain the results
- Handle missing values effectively
- Be capable of handling large datasets

Ensemble Learning



1980s — CART

1990s – 2000s — Bagging/Bootstrap Aggregating; Random Forests

2000s - 2020s — **Boosting Iteration**

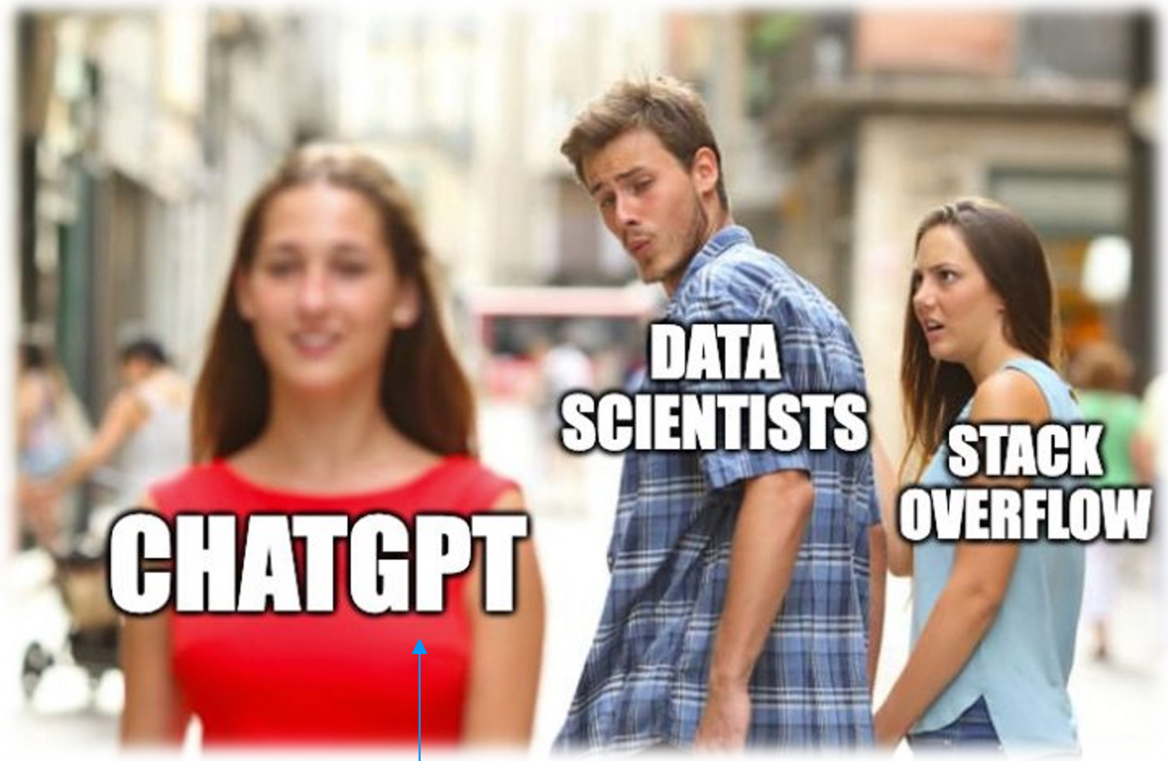
Boosting Iteration

1
1. Adaptive Boosting (AdaBoost)
2. AdaBoost Regression Task (AdaBoost.RT)
3. Log-Likelihood (LogitBoost)
4. Random Under Sampling (RUSBoost)
5. Synthetic Minority Over-Sampling Technique (SMOTEBoost)

2
6. **Gradient Boosting Machines (GBM)**
7. **Extreme Gradient Boosting (XGBoost)**
8. **Light GBM (lightGBM)**

3
9. Categorical Boosting (CatBoost)
10. Linear Programming Boosting (LPBoost)

Why Transformer Models?



Generative Pre-Trained Transformer (GPT)

- Cutting Edge Technique
- Parallelization
- Long-Range Dependencies
- Flexibility
- Pre-Training and Fine-Tuning
- State-of-the-Art Performance

GBM Model

Loss function $L(\theta)$ respect to coefficient θ

$$L(\theta) = \sum_{i=1}^{N} L[y_i, f(x_i;\theta)], \qquad \theta \in \mathbb{R}^p$$

$$\theta^* = arg\min_{\theta} L(\theta), \quad \theta^* = \sum_{b=0}^{B} \theta_b\,,$$

$$\{\theta_b\}_{b=1}^{B}$$

$$\theta_b = \theta_{b-1} + \gamma\Delta\theta_{b-1}, \qquad \gamma = \text{step size}$$

$$\Delta\theta_{b-1} = \left(-\frac{\partial L(\theta)}{\partial\theta}\right)$$

Loss function $L(f)$ respect to prediction function $\boldsymbol{f}$

$$L(f) = \sum_{i=1}^{N} L[y_i, f(x_i;\theta)],$$

$$f^* = arg\min_{f} L(f), \quad f_B = \sum_{b=0}^{B} f_b\ , f_b \in \mathbb{R}^N$$

$$f = \{f(x_i)\}_{i=1}^{N}$$

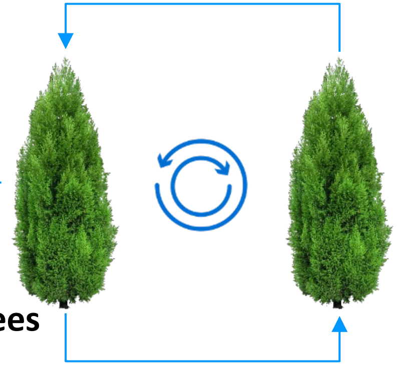
$$f_b = f_{b-1} - \gamma g_b, \qquad \gamma = \text{step size}$$

$$g_b = \left\{\left[\frac{\partial L(f)}{\partial f}\right]_{f=f_{b-1}(x_i)}\right\}_{i=1}^{N}$$

Boosting Tree

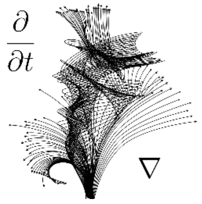
$$f^* = argmin_f L(f)$$

Minimize Errors $L(f)$



Weak Trees

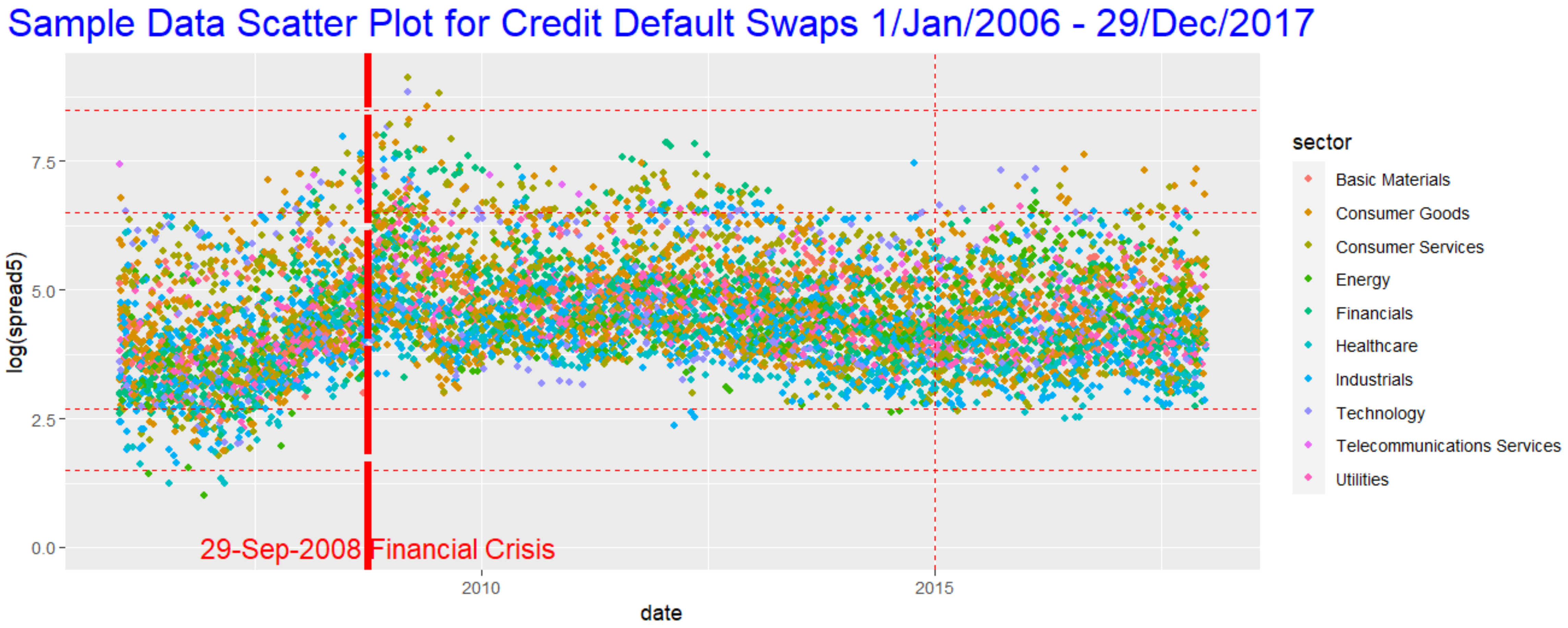
Training model
$f(x_i;\theta)$



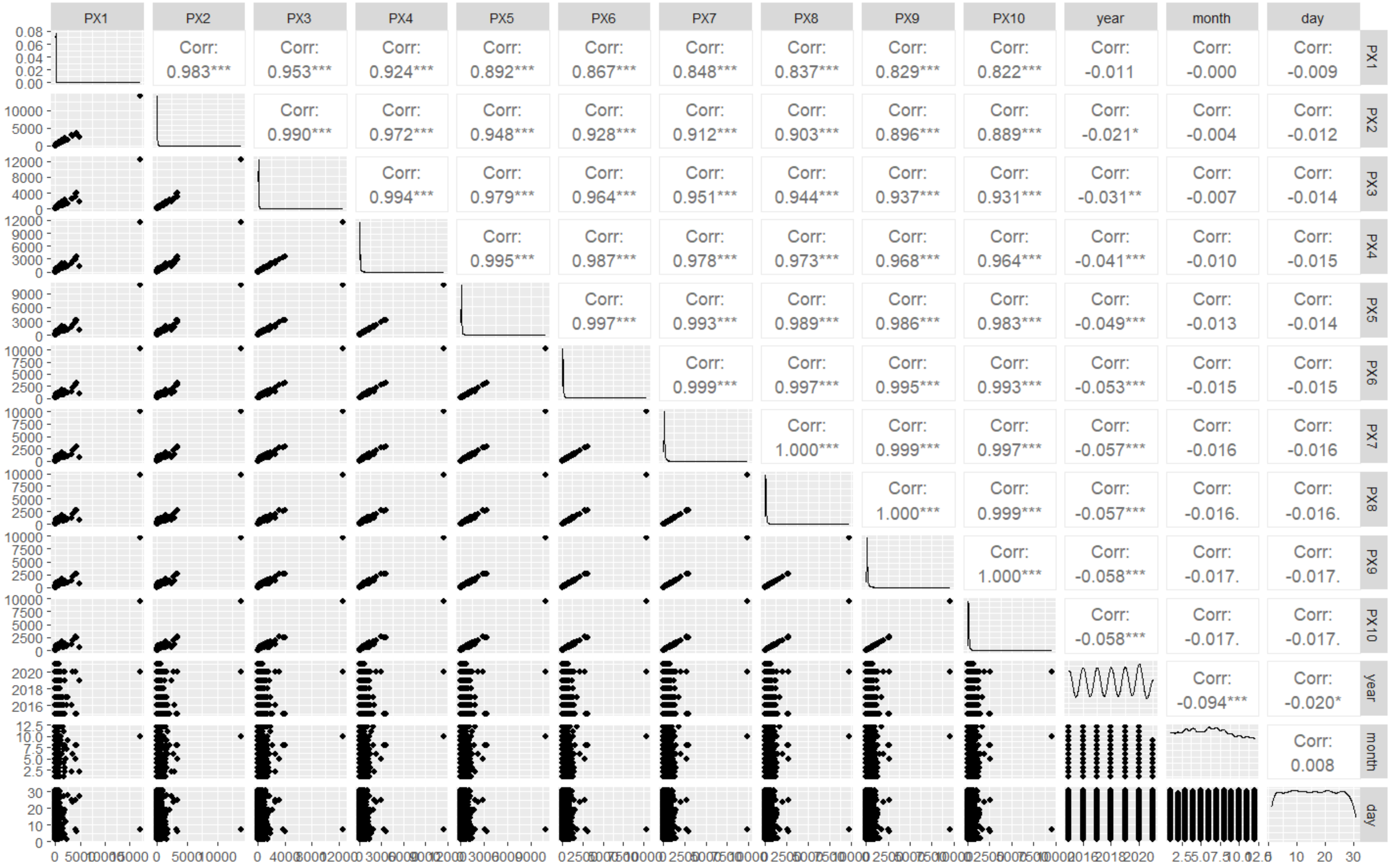
$$g_b = \left\{ \left[\frac{\partial L(f)}{\partial f} \right]_{f=f_{b-1}(x_i)} \right\}_{i=1}^{N}$$

What do we have?

A Bird's Eye View of Dataset From IMF

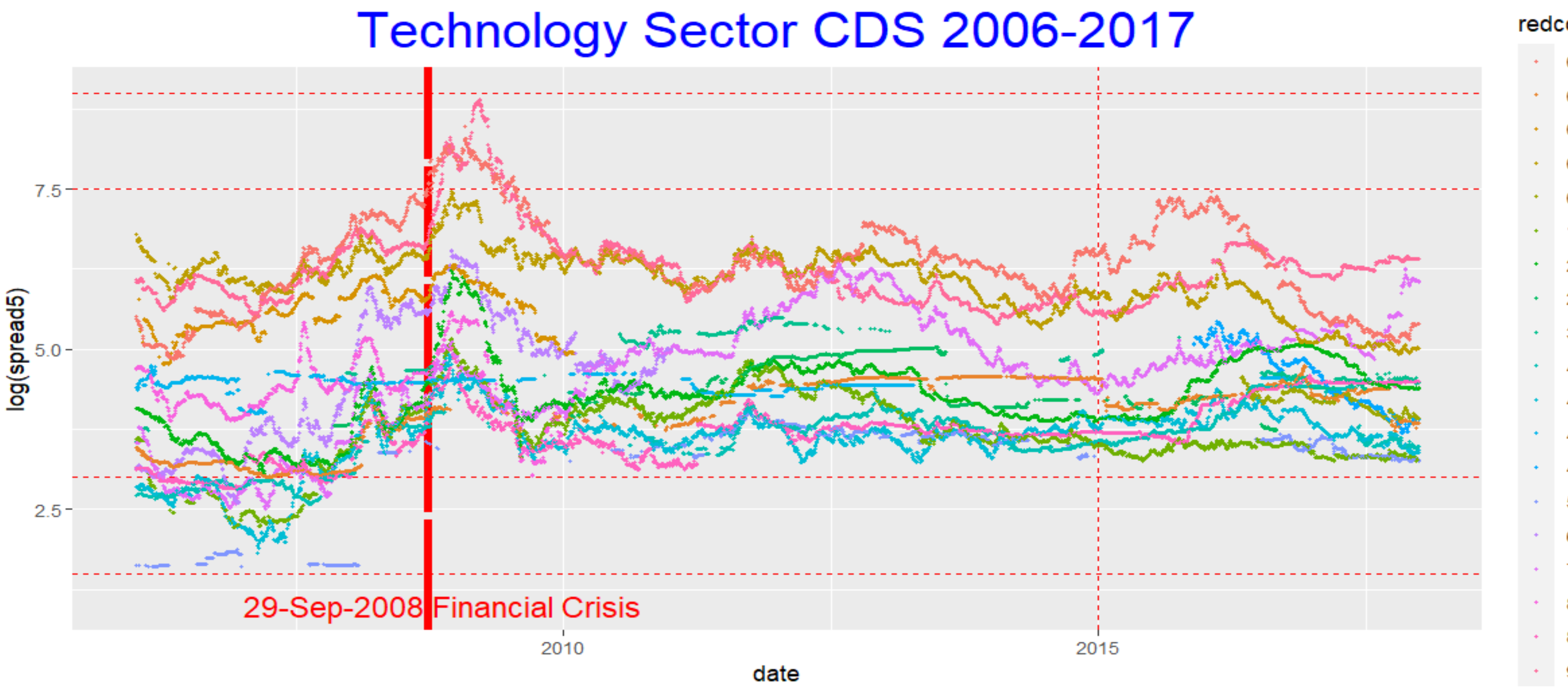


Different Types of CDS



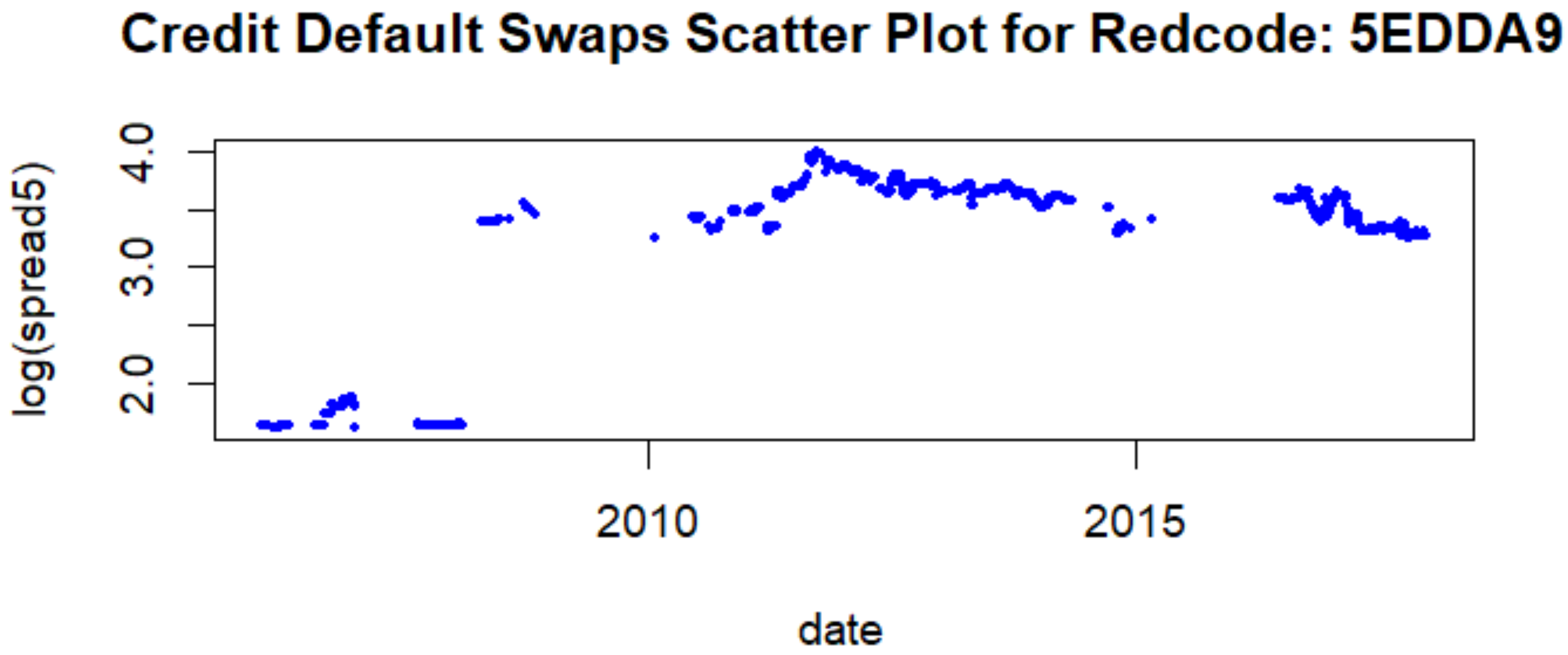
Sub-Dataset: Technology Sector

19 Companies Based on RED code



Sub-Dataset: Technology Sector

A particular Company

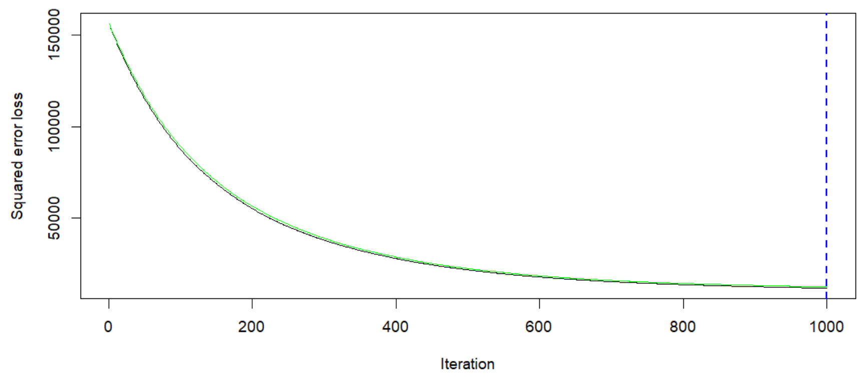


Experiment Setup

Sub-dataset = 37,526 observations, 139 features

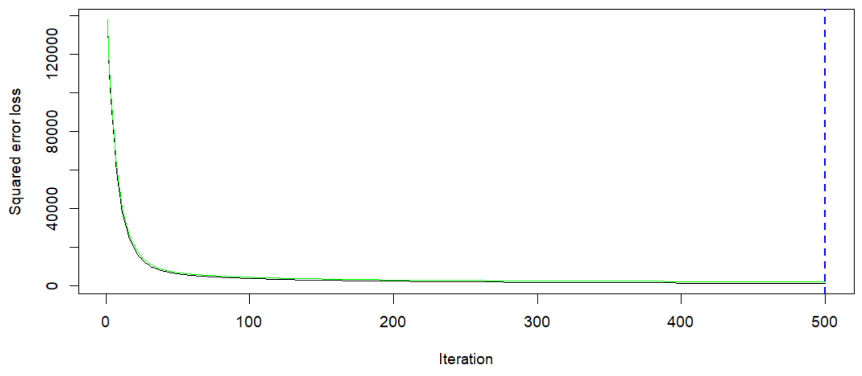
Increase:	Spread 5 By 10,000: common practice
Remove:	Dummy variables, left with 117 features
Split:	70:30 ratio, 70% for training, 30% for testing
CV fold:	5
HPC Config:	128 nodes + 256GB RAM

GBM Results



user **system** **elapsed**
36.00 0.47 99.82
gbm(formula = spread5 ~ ., distribution = "gaussian", data = cds_train,
n.trees = 1000, ***interaction.depth = 1***, **shrinkage = 0.01**, cv.folds = 5)
A gradient boosted model with gaussian loss function.
1000 iterations were performed.
The best cross-validation iteration was 1000.
There were 117 predictors of which 32 had non-zero influence.

RMSE = 112.5478



user **system** **elapsed**
50.30 0.56 147.31
gbm(formula = spread5 ~ ., distribution = "gaussian", data = cds_train,
n.trees = 500, **interaction.depth = 3**, **shrinkage = 0.1**, cv.folds = 5)
A gradient boosted model with gaussian loss function.
500 iterations were performed.
The best cross-validation iteration was 500.
There were 117 predictors of which 111 had non-zero influence.

RMSE =46.37817

The GBM Results Comparison

Parameters	Experiment 1	Experiment 2	Final Results
Distribution	Gaussian	Gaussian	Gaussian
# trees	1000	500	800
Shrinkage or learning rate	0.01	0.1	0.3
Interaction depth	1	3	5
# min. nodes	1	3	5
cv. fold	5	5	5
# predictors or features	117	117	117
Non-zero influence	32	111	**117**
Bag fraction	1	1	0.85
Train fraction	1	1	1
CPU usage time	36.00	50.30	111.31
System time	0.47	0.56	0.20
Elapsed time	99.82	147.31	132.05
RMSE	112.548	46.372	29.512

Is it optimal?

Xgbm Run 243 Grid Points Hyperparameter Search on HPC

Learning rate: (0.1, 0.3, 0.5)
Depth of trees: max # of tree depth (5,7,9)
Min child weight: min # of observation in each terminal node (3,5,7)
Subsample: controls a fraction of the training observation (0.65, 0.8, 1)
Column Sample: percentage of columns (0.65, 0.8, 1)

$$\boldsymbol{grid\ points} = \begin{bmatrix} 0.10 & 0.30 & 0.50 \\ 5 & 7 & 9 \\ 3 & 5 & 7 \\ 0.65 & 0.80 & 1 \\ 0.65 & 0.80 & 1 \end{bmatrix}$$

$$= \mathbf{3^5} = \mathbf{243}$$

Parameters	CPU usage time	System time	Elapsed time
HPC platform (sec.)	**593,717.98**	**69.55**	**4,713.02**
Shrinkage or learning rate	Max tree depth	Min. rows /each end node	k fold CV
0.10	9	1	5
Subsample for each tree	Column sample	Number of trees	Min RMSE
0.80	1	250	25.70

Total Trees = 250 X 243 = 60,750

Can We Trust Result?

$$[\forall \boldsymbol{C}]\exists[\boldsymbol{R}, \boldsymbol{D}, \boldsymbol{O}] \vDash (\exists\ \boldsymbol{P})$$

$\boldsymbol{R}$: $f(x) = g(x_1, x_2, \cdots, x_n)$

$$\boldsymbol{f} = \{\boldsymbol{f}(\boldsymbol{x}_i)\}_{\boldsymbol{i}=\boldsymbol{1}}^{\boldsymbol{N}}$$

$\boldsymbol{O}$: $\theta^* = \underset{\theta}{argmin}\, L[y_i, f(x_i; \theta)]$

$$\boldsymbol{g_b} = \left\{\left[\frac{\boldsymbol{\partial L(f)}}{\boldsymbol{\partial f}}\right]_{\boldsymbol{f}=\boldsymbol{f_{b-1}(x_i)}}\right\}_{\boldsymbol{i}=\boldsymbol{1}}^{\boldsymbol{N}} \longrightarrow \textbf{convergence}$$

Creation (Models)
- **Explainability/Interpretability**
- Transparency
- Justice
- Fairness

Evaluation (Data)
- Availability
- Usability
- Security/Privacy
- Accountability

Selection (Algorithms)
- Robustness
- Reproducibility
- Reliability
- Accuracy

RMSE or MSE: $\boldsymbol{D}$: $L(f) = \sum_{i=1}^{N} L[y_i, f(x_i; \theta)] \longrightarrow$ **Data Governance**

What is explanation / interpretation?

UNIVERSITÉ DU LUXEMBOURG

Reason

Comply with Law of Nature

Comply with Law of Heart

Explanation

- Cause effect
- Induction
- Deduction
- Abduction
- Meta-Reason/Framework
- Counterfactual

Interpretation

- Ethics/Morality
- Virtue/Integrity
- Justice
- Fairness
- Value/Principle/Motivation
- Emotion/Spirit/Culture

Reason

Reason for Reason itself

Outward Reasoning

Reasons want to make sense

- Logic
- Rationality
- Inference
- Assumption
- Hypothesis

Inward Reasoning

Reasons want to be happy

- Freedom
- Passion
- Choices
- Belief/Trust
- Perception

Many Ways to Explain
From Decision-Making Perspective

Strategic

1. **Feature Importance: VI**
2. **Global Explanations: PDP, ICE**
3. Counterfactual Explanation
4. Meta-Explanations (Ensemble)
5. Causal Explanations
6. Integrated Gradients
7. Graph-based Explanations
8. Concept-Based Explanation

Tactical

1. Prototype-based Explanation
2. Confidence Intervals
3. Model-Agnostic Explanation
4. Surrogate Model
5. Certified Explanations
6. Rule-Based Explanation
7. Layer-wise relevance propagation (LRP)
8. Model Debugging

Operational

1. **Local explanations LIME, SHAP**
2. Instance-Based Explanation
3. Sensitivity Analysis
4. Simulatability
5. Behavioural testing
6. Activation Maximization
7. Interactive Dashboards
8. Attention mechanisms

TAI: We Use Five Techniques

Global

- Variable Importance (VI)
- Partial Dependent Plot (PDP)
- Individual Conditional Expectation (ICE)

Local

- Local Interpretable Model-agnostic Explanations (LIME)
- Shapley Values (SHAP)

Math Expressions of All Explanatory Techniques

- **Feature importance (VI)**

$$MDA(X_i) = \frac{1}{n}\sum_{j-1}^{n}\left(f(X) - f(X_{ij})\right)$$

n is the number of permutations, X is the original dataset
X_{ij} is the dataset with the i-th feature values permutation in the j-th permutation

- **PDP**

$$\hat{f}_s(x_s) = E_{X_c}[\hat{f}_s(x_s, X_c)] = \int \hat{f}_s(x_s, X_c)\, d\mathbb{P}(X_c)$$

$$\hat{f}_s(x_s) = \frac{1}{n}\sum_{i=1}^{n}\hat{f}\left(x_s, x_c^{(i)}\right)$$

x_s is the feature , X_c other features

- **ICE (Individual Conditional Expectation)**

$$\hat{f}_S^{(i)} = \left\{x_S^{(i)}, x_c^{(i)}\right\}_{i=1}^{N} ; \quad \hat{f}_{cent}^{(i)} = \hat{f}^{(i)} - 1\hat{f}\left(x^a, x_c^{(i)}\right)$$

x^a is the anchor point; $\hat{f}$ is fitting model

LIME

$$explain(x) = arg \min_{g \in G}[L(f, g, \pi_x) + \Omega(g)]$$

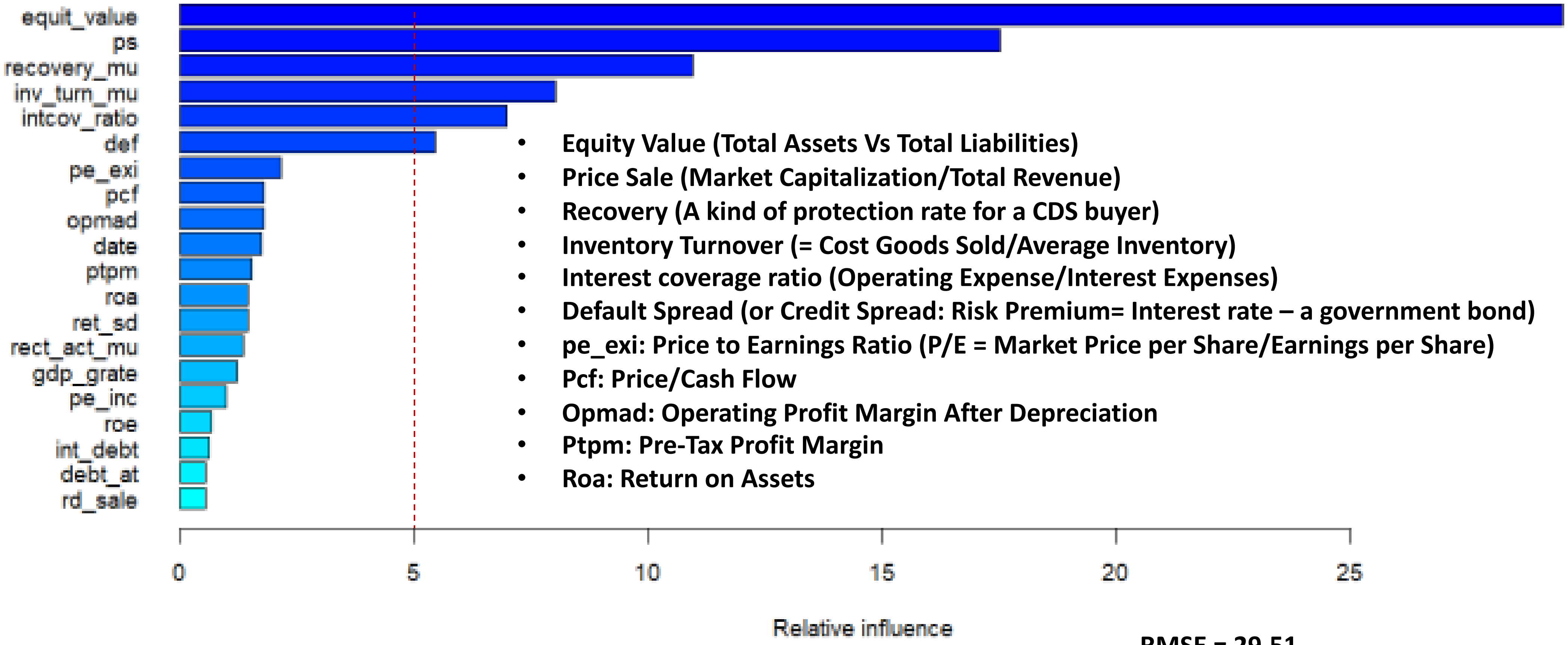
x = instance
g = the model (e.g., linear regression model)
L= loss function
f = predictive model
$\Omega(g)$= a model complexity

- **SHAP**

$$g(z') = \phi_0 + \sum_{j=1}^{M}\phi_j z'_j, \qquad z' \in \{0,1\}^M$$

g= explanation model
$z' \in \{0,1\}^M$ is the coalition vector
M is the max coalition size
$\phi_j \in \mathbb{R}$ is the feature attribution for a feature j

Variable Importance

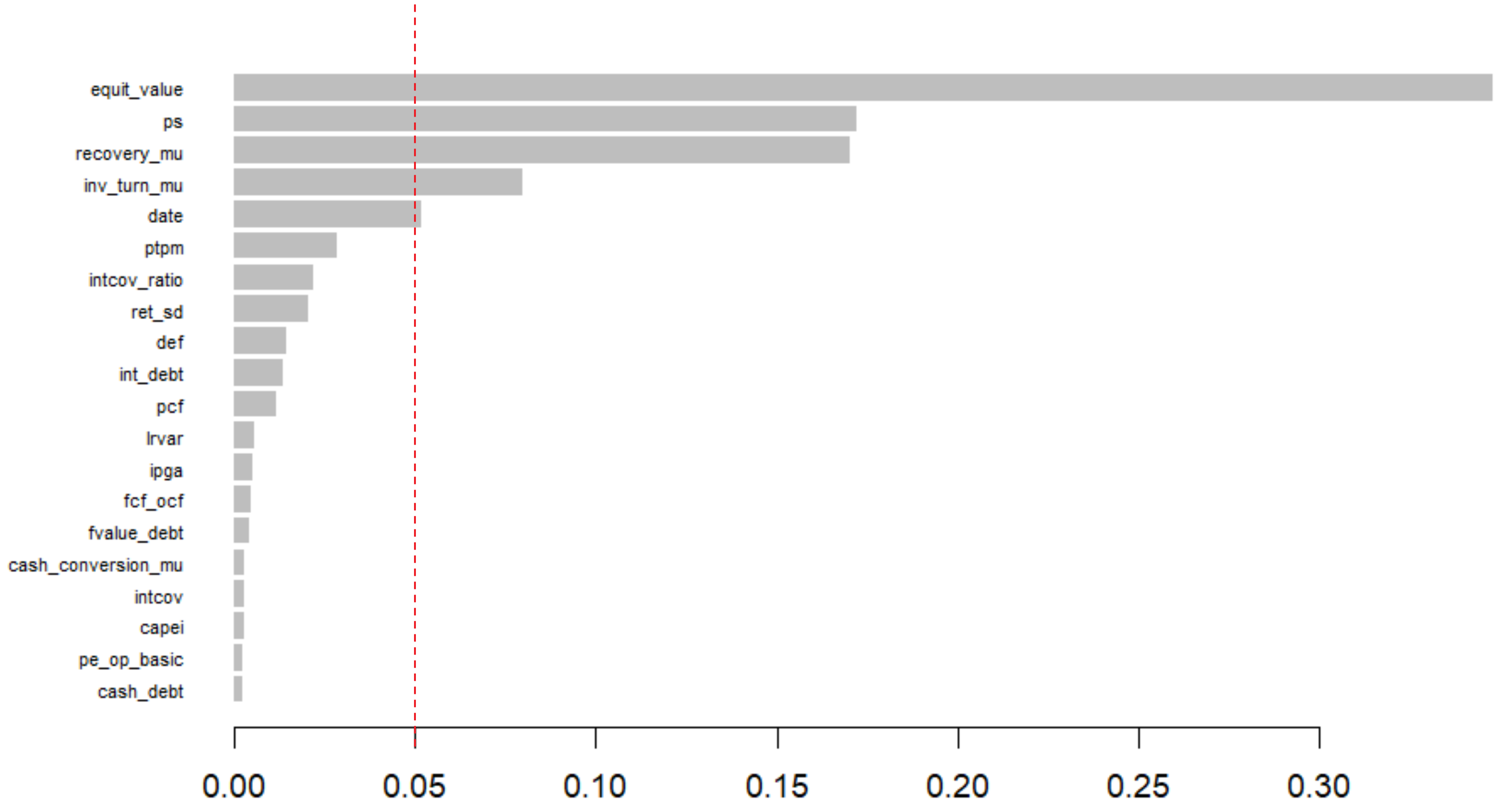


- **Equity Value (Total Assets Vs Total Liabilities)**
- **Price Sale (Market Capitalization/Total Revenue)**
- **Recovery (A kind of protection rate for a CDS buyer)**
- **Inventory Turnover (= Cost Goods Sold/Average Inventory)**
- **Interest coverage ratio (Operating Expense/Interest Expenses)**
- **Default Spread (or Credit Spread: Risk Premium= Interest rate – a government bond)**
- **pe_exi: Price to Earnings Ratio (P/E = Market Price per Share/Earnings per Share)**
- **Pcf: Price/Cash Flow**
- **Opmad: Operating Profit Margin After Depreciation**
- **Ptpm: Pre-Tax Profit Margin**
- **Roa: Return on Assets**

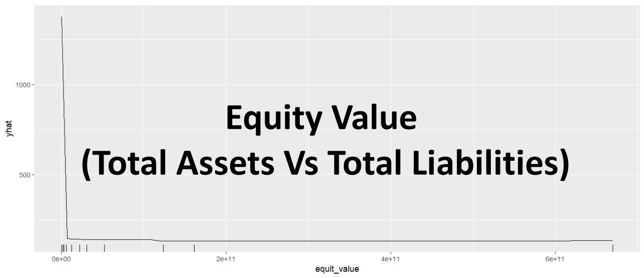
RMSE = 29.51

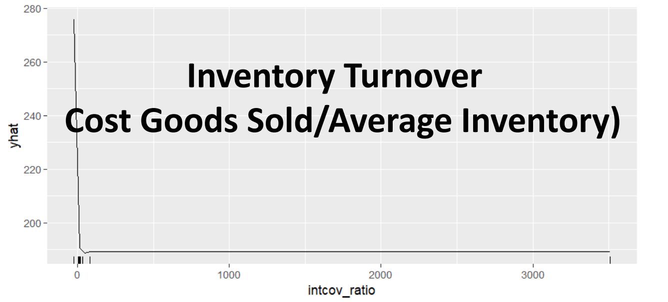
Variable Importance

RMSE = 25.70

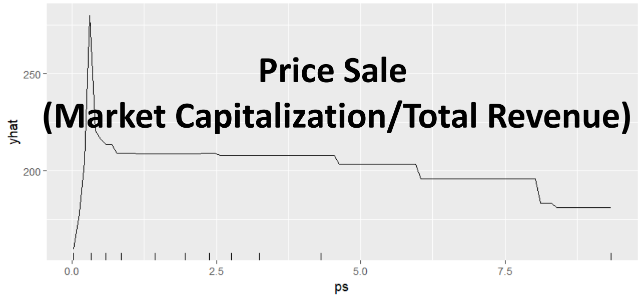


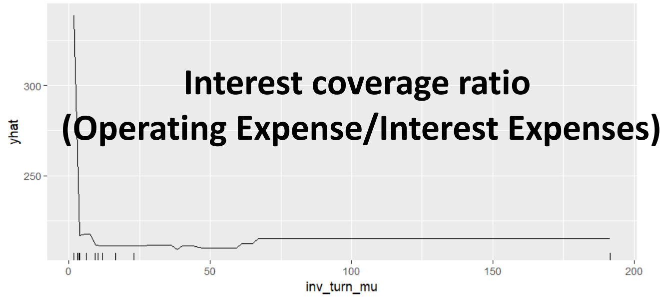
PDP

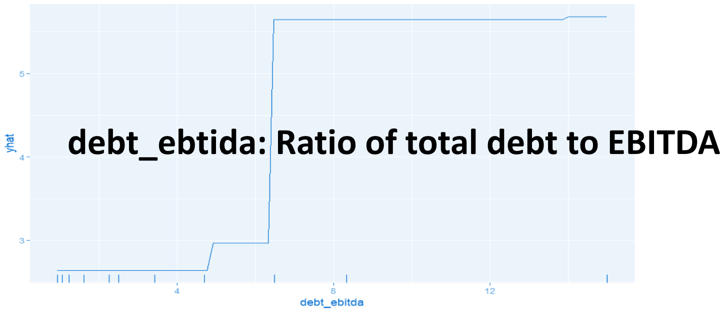

Equity Value
(Total Assets Vs Total Liabilities)
yhat
equit_value

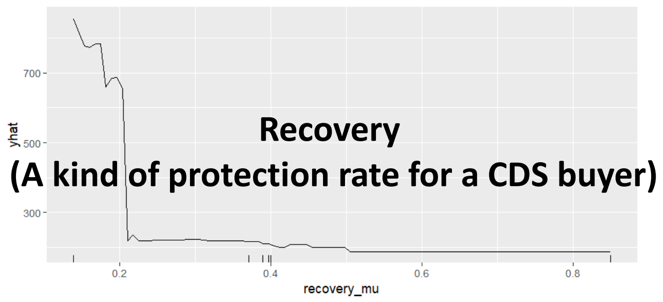

Inventory Turnover
Cost Goods Sold/Average Inventory)
yhat
intcov_ratio

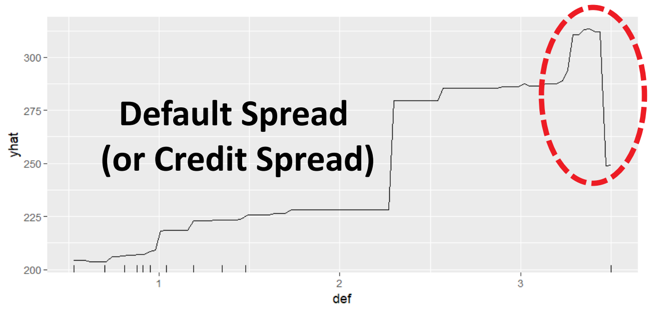

pe_exi: Price to Earnings Ratio
yhat
pe_exi

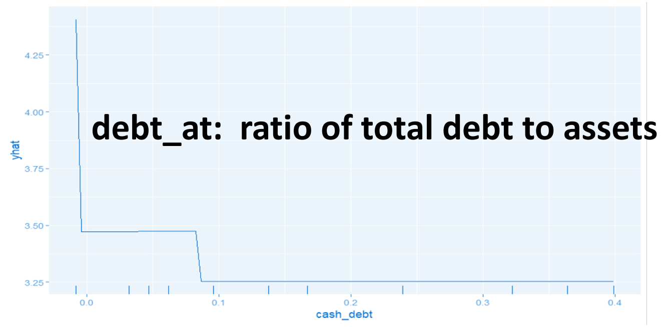

Price Sale
(Market Capitalization/Total Revenue)
yhat
ps


Interest coverage ratio
(Operating Expense/Interest Expenses)
yhat
inv_turn_mu

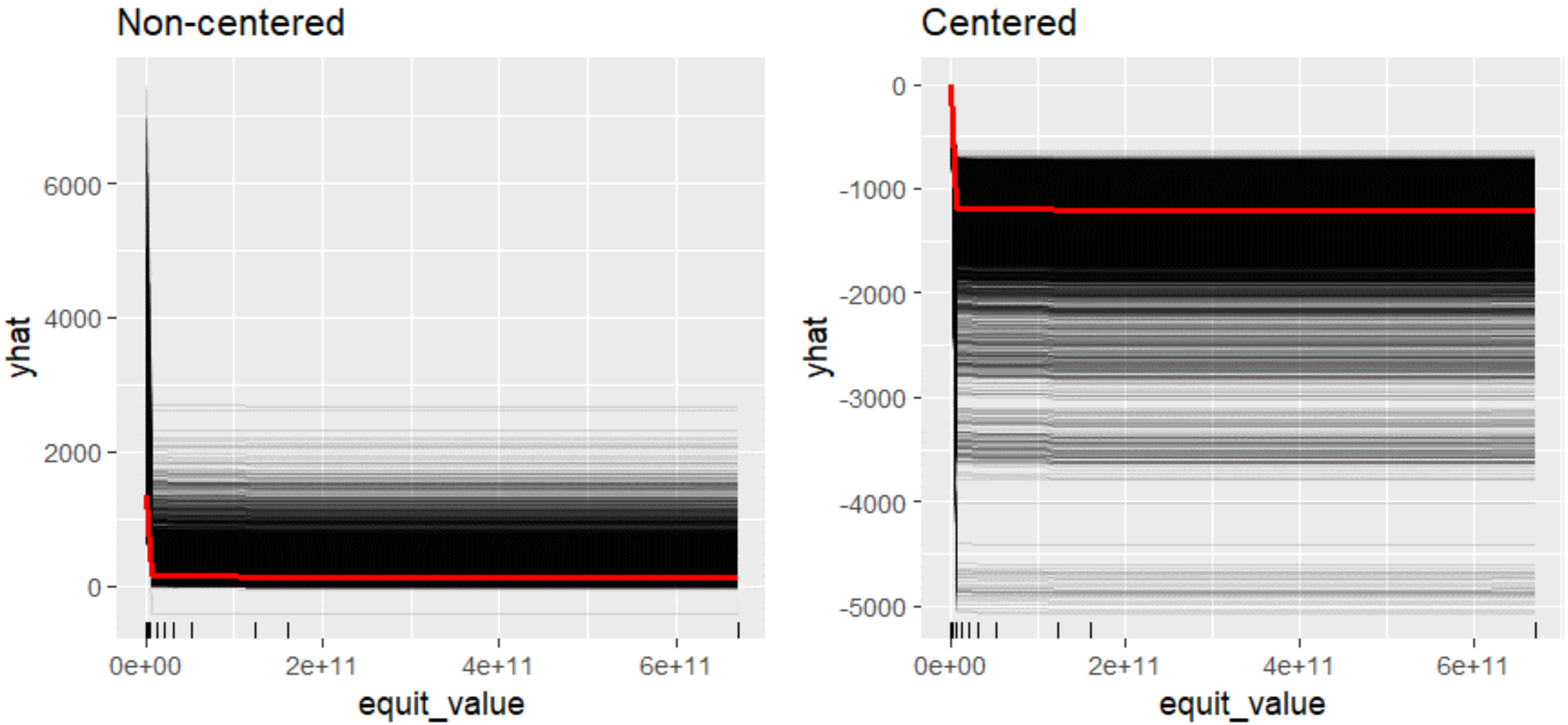

debt_ebtida: Ratio of total debt to EBITDA
yhat
debt_ebitda


Recovery
(A kind of protection rate for a CDS buyer)
yhat
recovery_mu


Default Spread
(or Credit Spread)
yhat
def

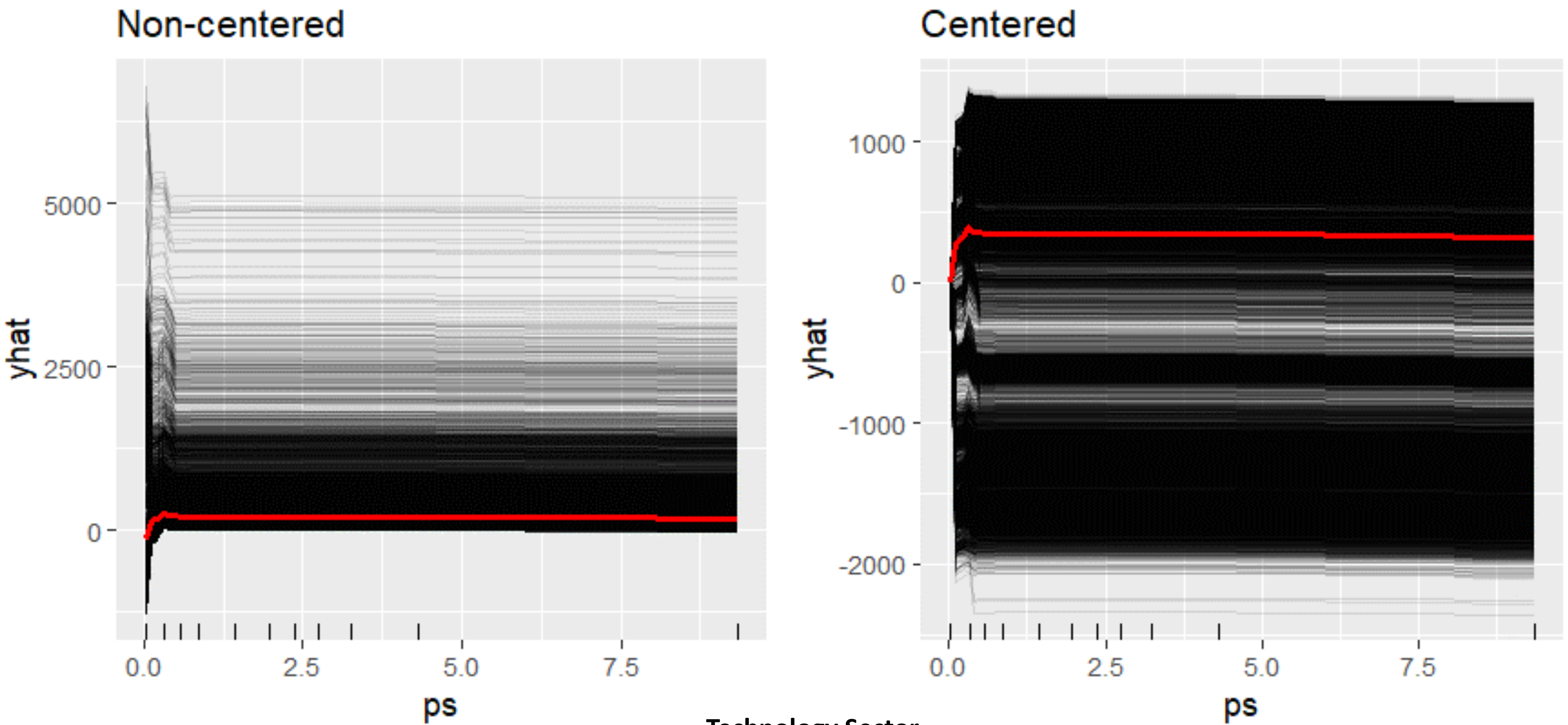

debt_at: ratio of total debt to assets
yhat
cash_debt

ICE of Equity Value



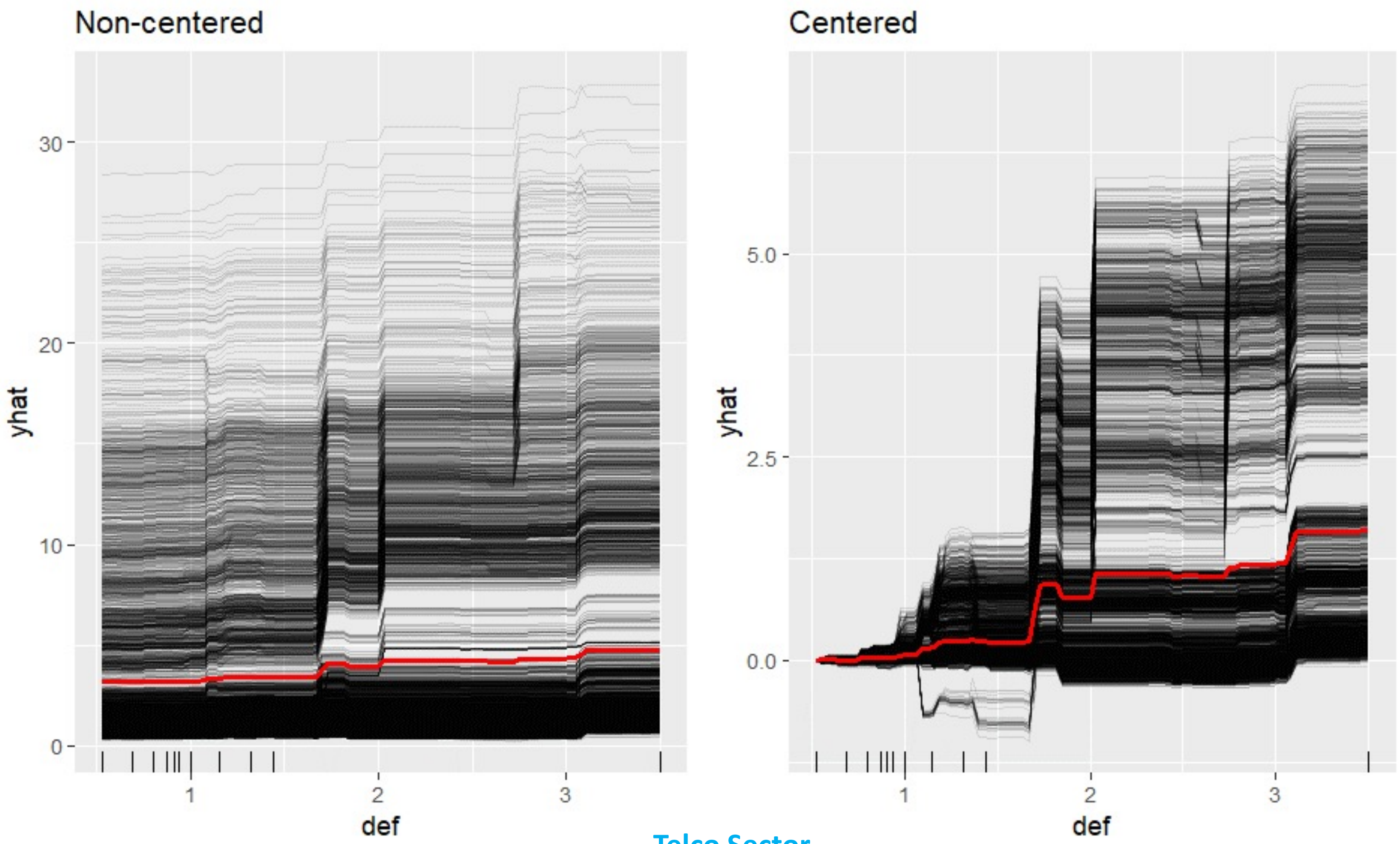
Technology Sector

ICE of Price Sale



Technology Sector

ICE of Default

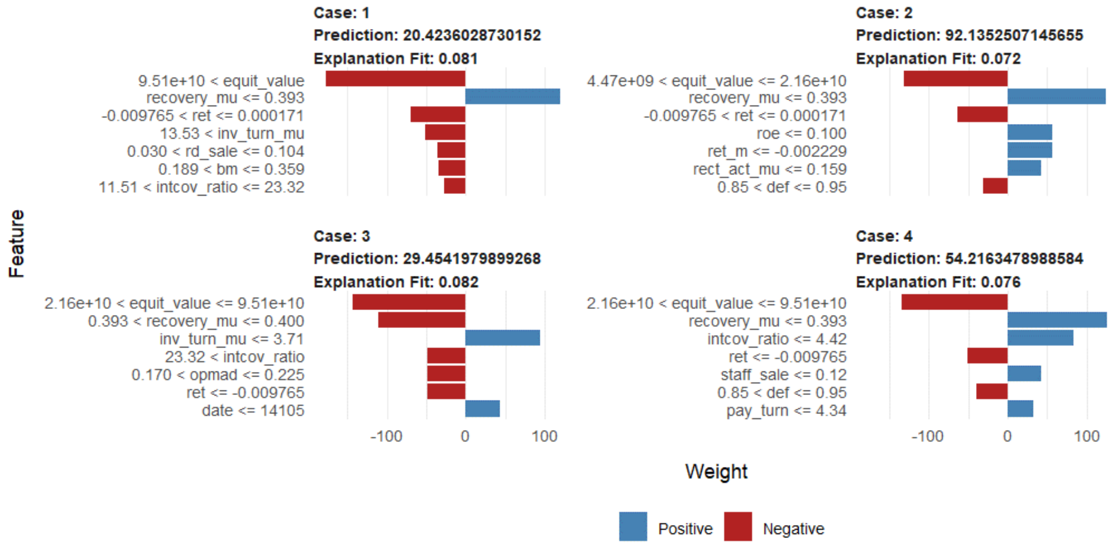


Telco Sector

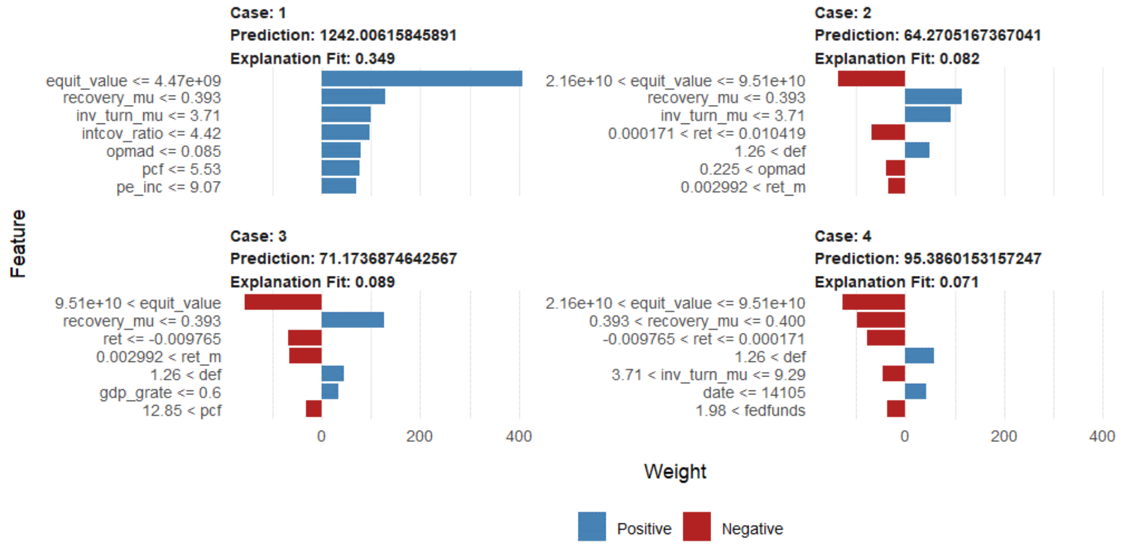
UNIVERSITÉ DU LUXEMBOURG

LIME

Before 2008 Financial Crisis

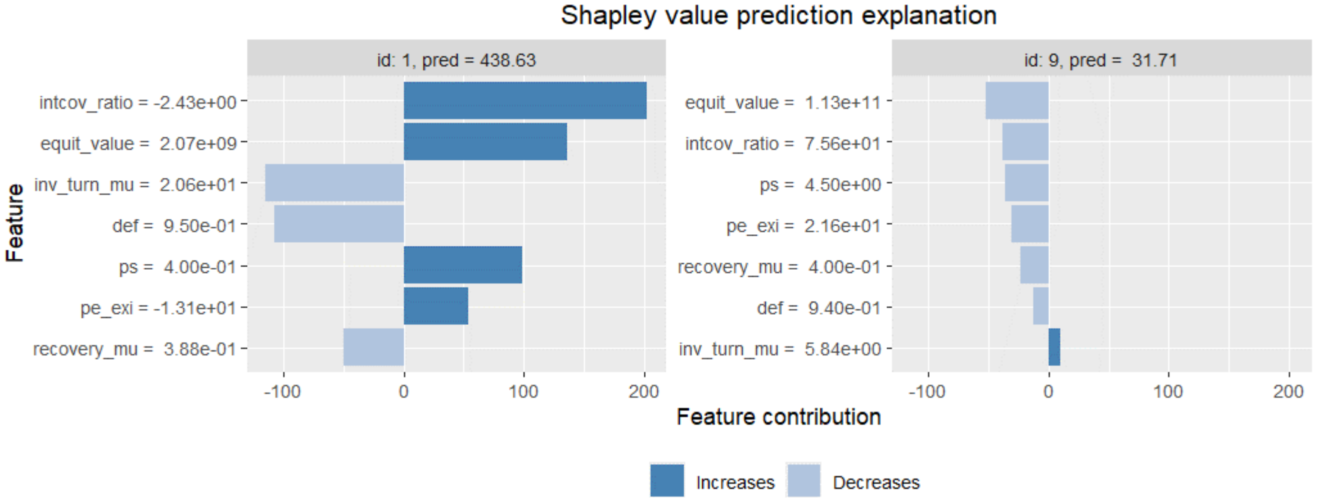


After 2008 Financial Crisis

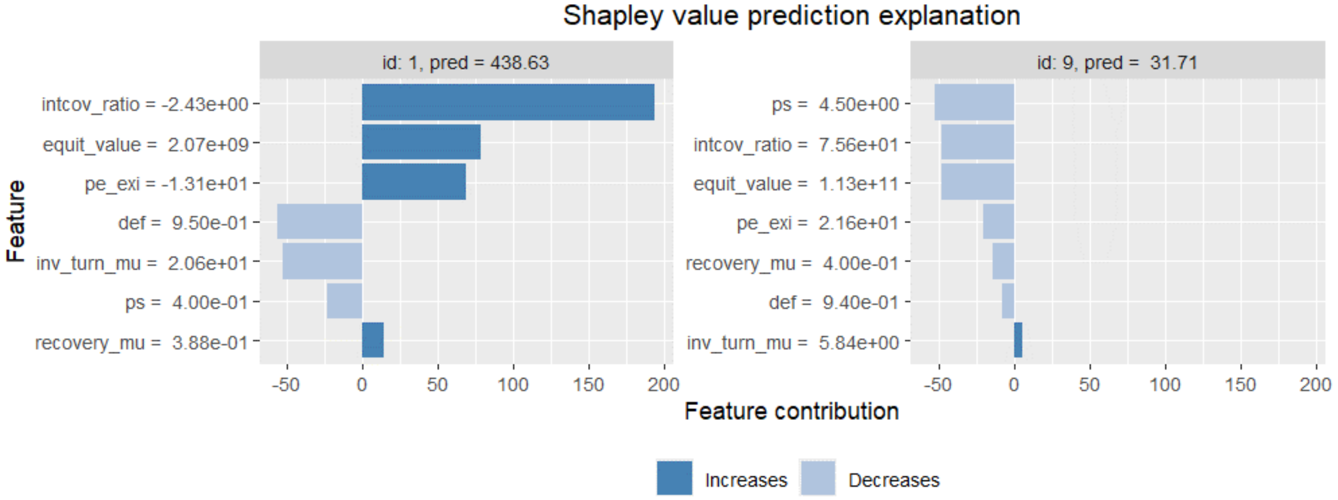


SHAP

Empirical



Copula



Conclusion

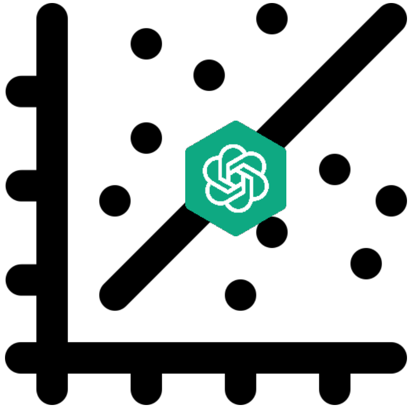
We can trust AI by the law of nature,
Can we trust AI by the law of heart?

Challenging!

Is a reason reasonable? :

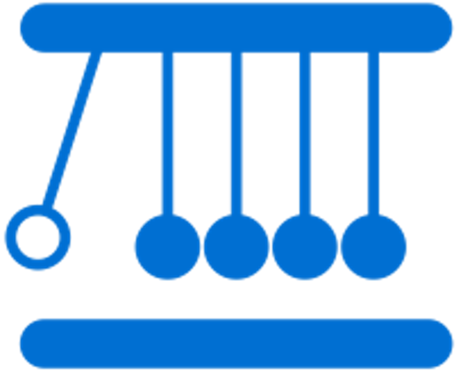
$$\boldsymbol{P \vDash \pi_A \vDash \pi_B \vDash \pi_C, \cdots,}$$

What is the issue of AI/ML?



Correlation

≠



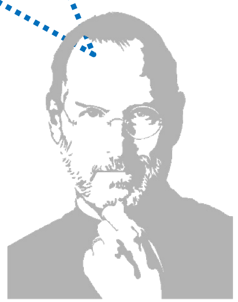
Causation

However, There is A Catch

"You cannot connect the dots looking forward; you can only connect them looking backwards."

Strategic decision-making requires placing dots by looking forward

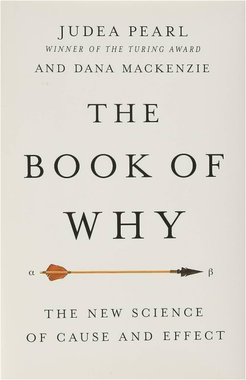
AI/ML is the same as connecting dots

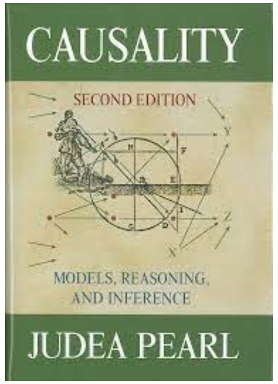


Paradox or Dilemma?

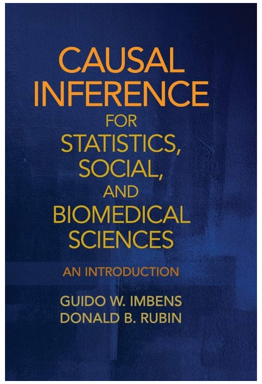
UNIVERSITÉ DU LUXEMBOURG

Future Research: Causal Inference



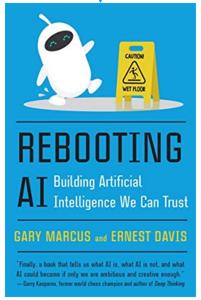


Judea Pearl: Turing Award winner 2011



Guido Imbens: Nobel Laureate 2021

Doug Lenat Gary Marcus



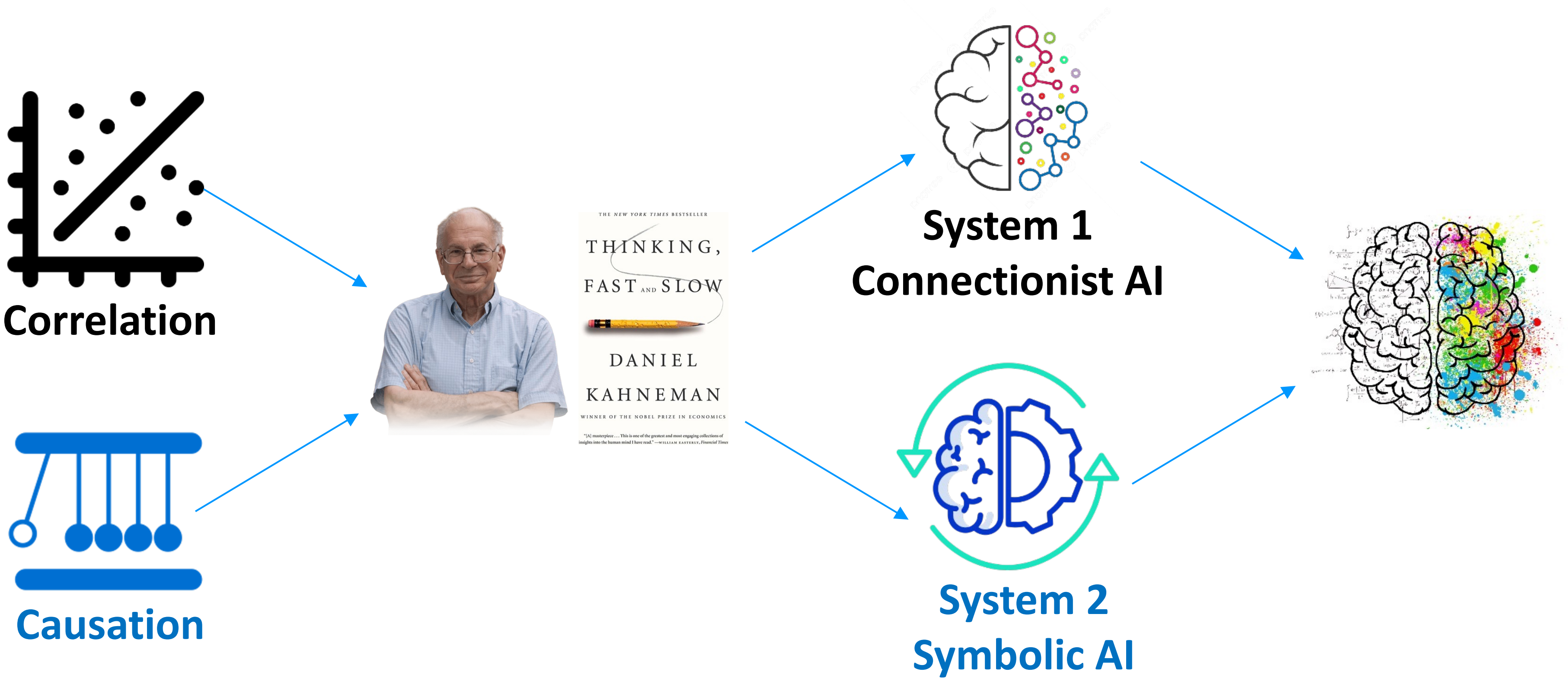
Rebooting AI: Building AI We Can Trust

Commonsense knowledge inference

Recent Paper 31/Aug/2023

“Getting from Generative AI to Trustworthy AI: What LLMs might learn from Cyc”

Future Research Direction



Q&A