Exploring Machine Learning

In the context of risk modelling of home loans

Andreas, Rana, Robin

ETH zürich | CREDIT SUISSE

Executive Summary

QUESTION

How can Credit Suisse leverage **Machine Learning** in prediction of **Probability Of Default** when evaluating loan applications?

CONSIDER

Bias	Explainability	Accuracy

STRATEGY

Delivering new insights using Machine Learning

Counterfactual Analysis	Neural Networks

SUGGEST

Automation & Assistance

Machine learning as a guiding tool for credit officers to help them determine accurate judgement alongside online tools for individuals to understand their credit health and chances at getting a successful application.

Introducing Roger

Roger
35

Value

Considering expanding his organic café to a second location. Faces the challenge of securing a loan.

Challenge

His first application was denied, leaving him confused and disheartened.

Reasoning

The feedback was vague, citing his "risk profile" as a concern, but with no clear understanding of what that meant.

There is an **opportunity** for Credit Suisse to pave the way for understandable credit options.

Introducing Roger

Value

loan.

Challenge

His first application was denied, leaving him confused and disheartened.

understanding of what that meant.

What should it look like for Roger?

There is an opportunity for Credit Suisse to pave the way for understandable credit options.

The Credit Risk industry has built itself on stability and simplicity

Security

Stability

Prudent

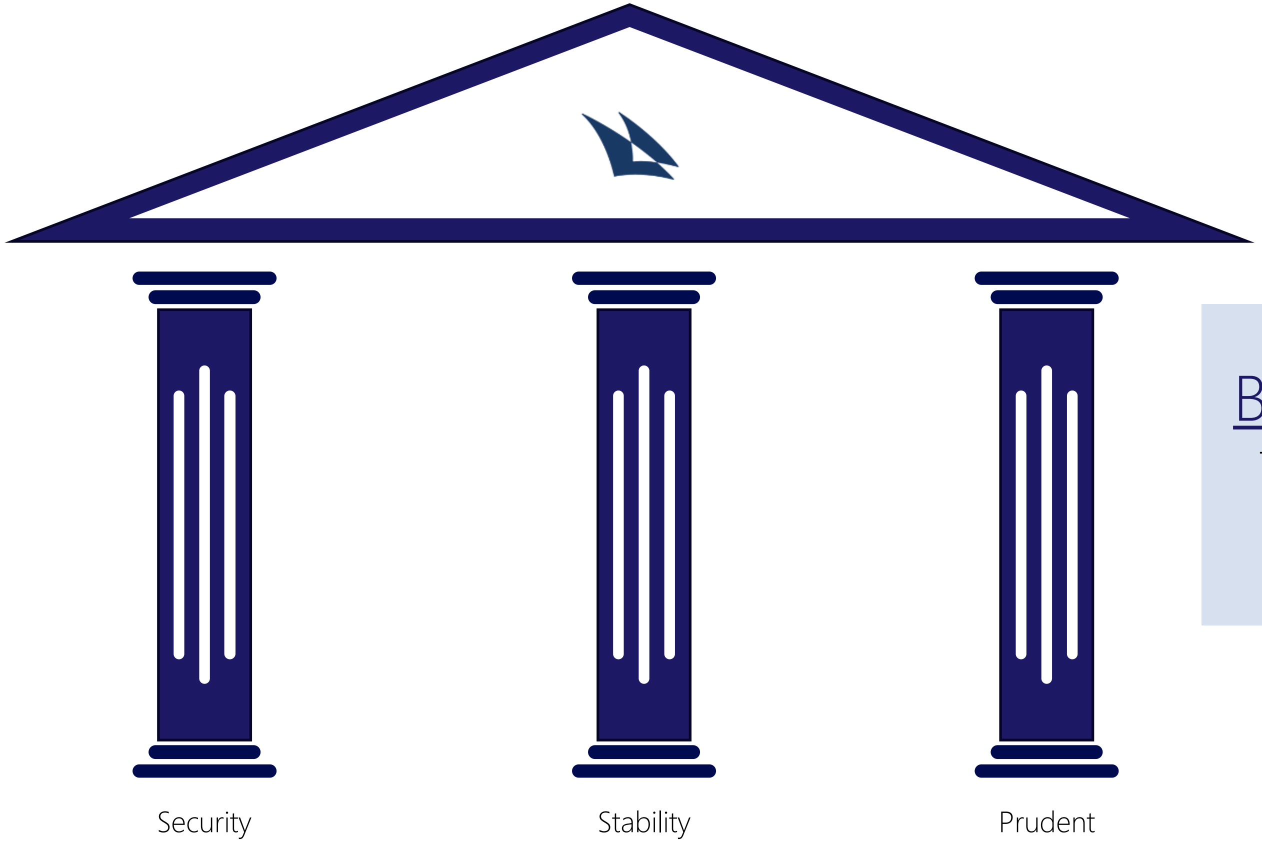
Why should we change what isn't broken?

Credit officers trust their clients

Trust is a key cornerstone of the relationship between Credit Suisse and their clients.

Attention to Detail

More accurate, efficient and personalized risk assessments.



Bridging tradition with innovation

The melding of Swiss banking's privacy and trust with machine learning for precise, reliable home loan risk assessments

Where does advanced computing fit within risk modelling?

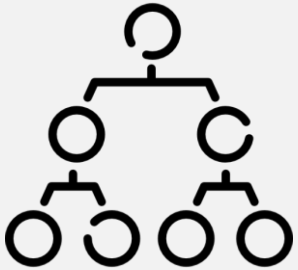
Current

Advanced computing

Statistical Learning

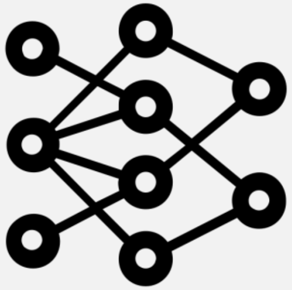
Linear Regression
Linear Discriminant Analysis

Machine learning



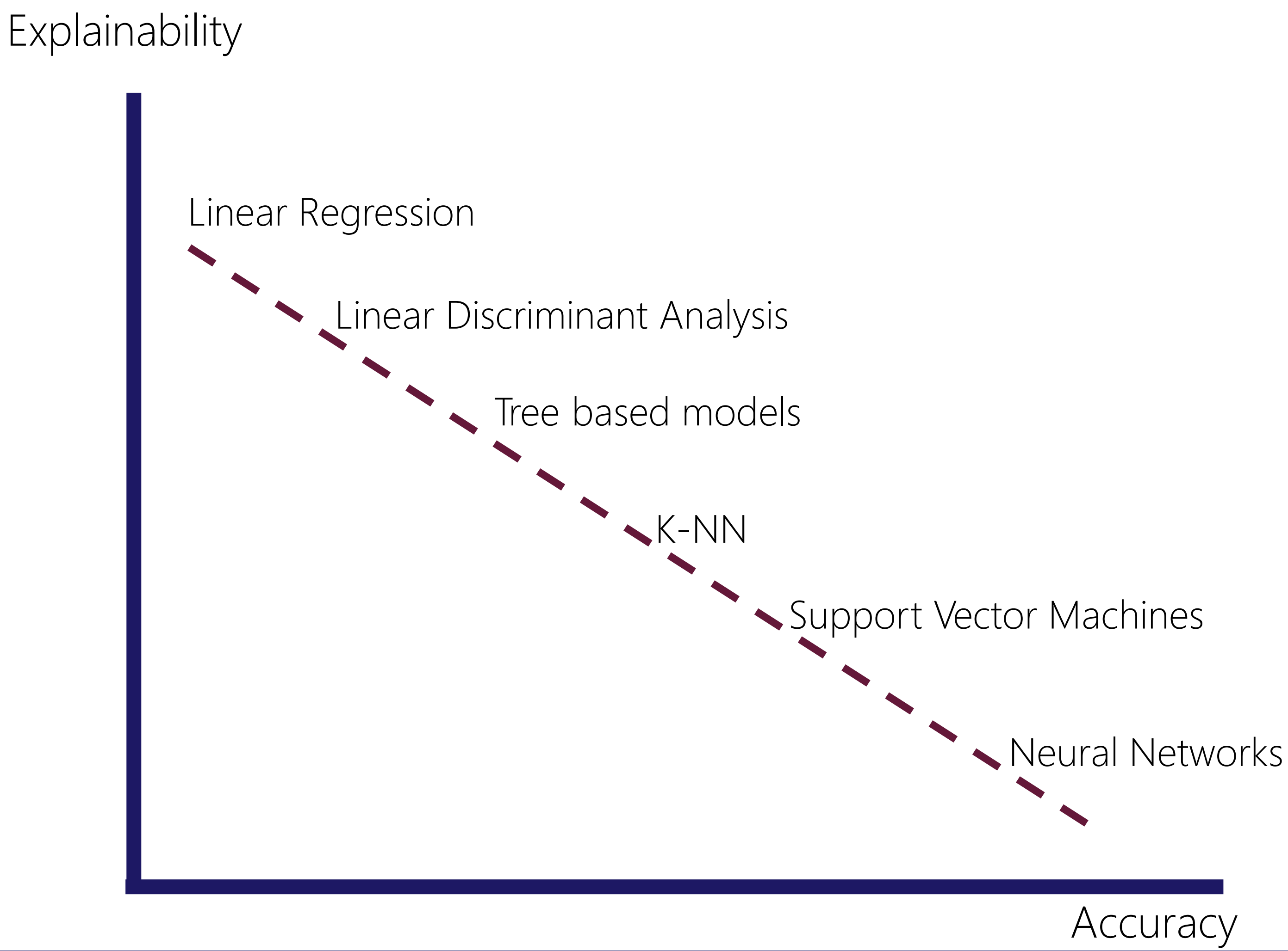
K-NN
Tree based models
Support Vector Machines

Deep learning



Neural Networks

Where does advanced computing fit within risk modelling?



Black box model
Models which utilize complex decision criteria, making predictions without clear explanations

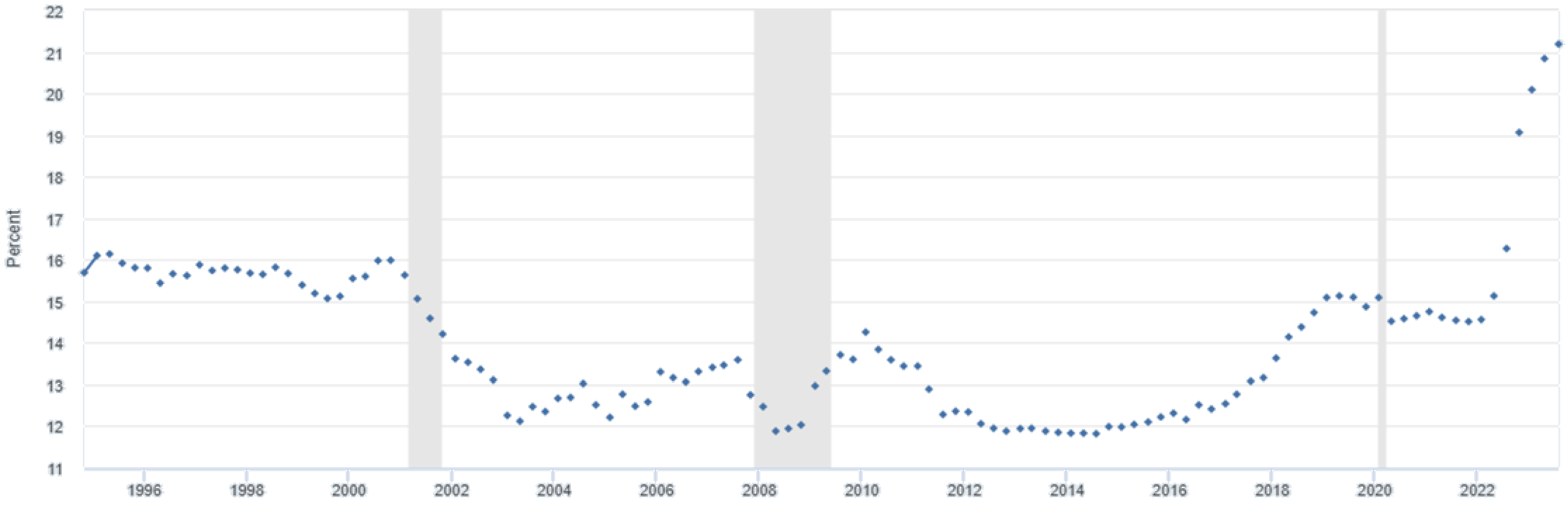
Transparent model
Decision making process is grounded in logic and easily interpreted by humans.

Is this necessarily the case here?

Institutions are hesitant to go all in with machine learning despite claims

Goldman Sachs

"even when credit scoring is done in compliance with the law, it can reflect and perpetuate societal inequality."

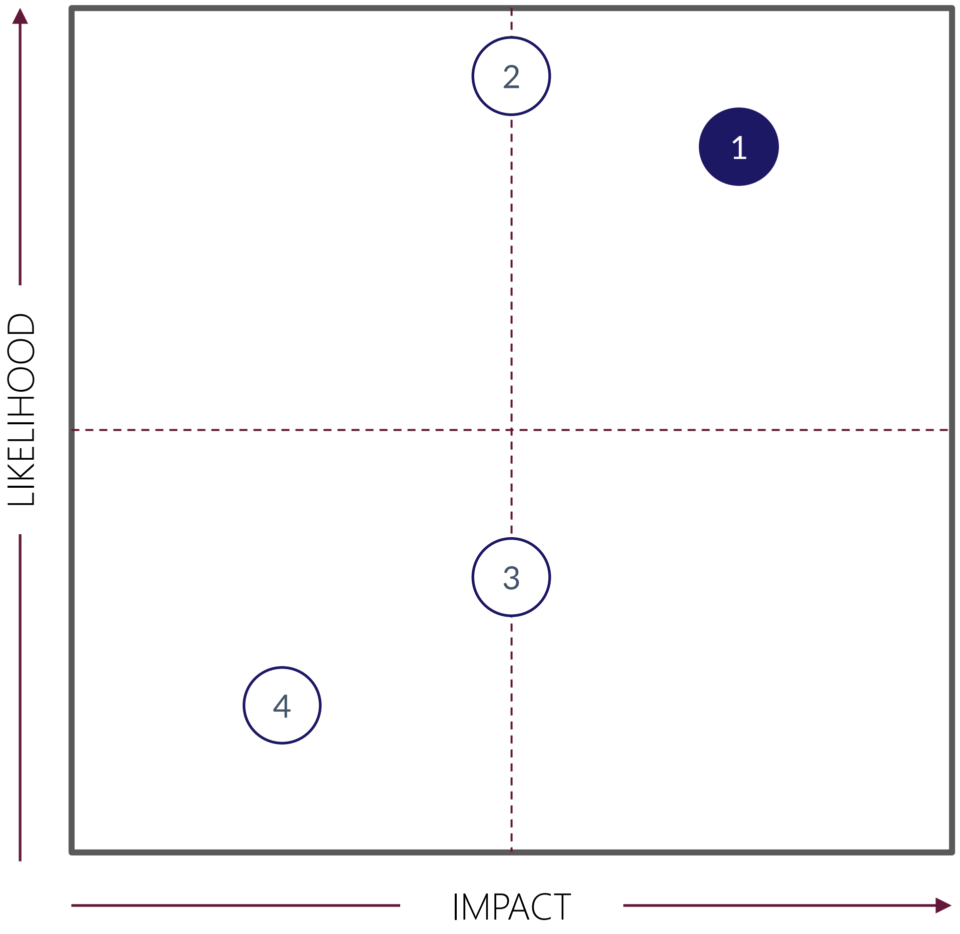


Despite the advances in Machine Learning CC rates are at all-time highs

US Commercial bank CC interest rates

Key Concerns When Implementing Machine Learning Into Risk Modelling

	Features	Concerns
1	Bias	Bias introduced as a result of proxy variables and non-inclusive data ranges.
2	Transparency	Methods must be explainable to relevant stakeholders and committees.
3	Regulations / Privacy	Fair lending laws, data privacy violations as a result of machine learning.
4	Substantial benefit	Overcomplicated model creation without any significant increase in predictive capabilities



The Focus of our research was within the Modelling stage.

Preprocessing

Filtering

Interactions

Transforming and normalizing data for optimal model performance.

Modelling

Precision

Explainability

Applying algorithms to create predictive or analytical models.

Tuning

Parameters

Validation

Adjusting model parameters to enhance accuracy and efficiency

The Focus of our research was within the Modelling stage.

Preprocessing

Filtering

Interactions

Transforming and normalizing data for optimal model performance.

Modelling

Precision

Explainability

Applying algorithms to create predictive or analytical models.

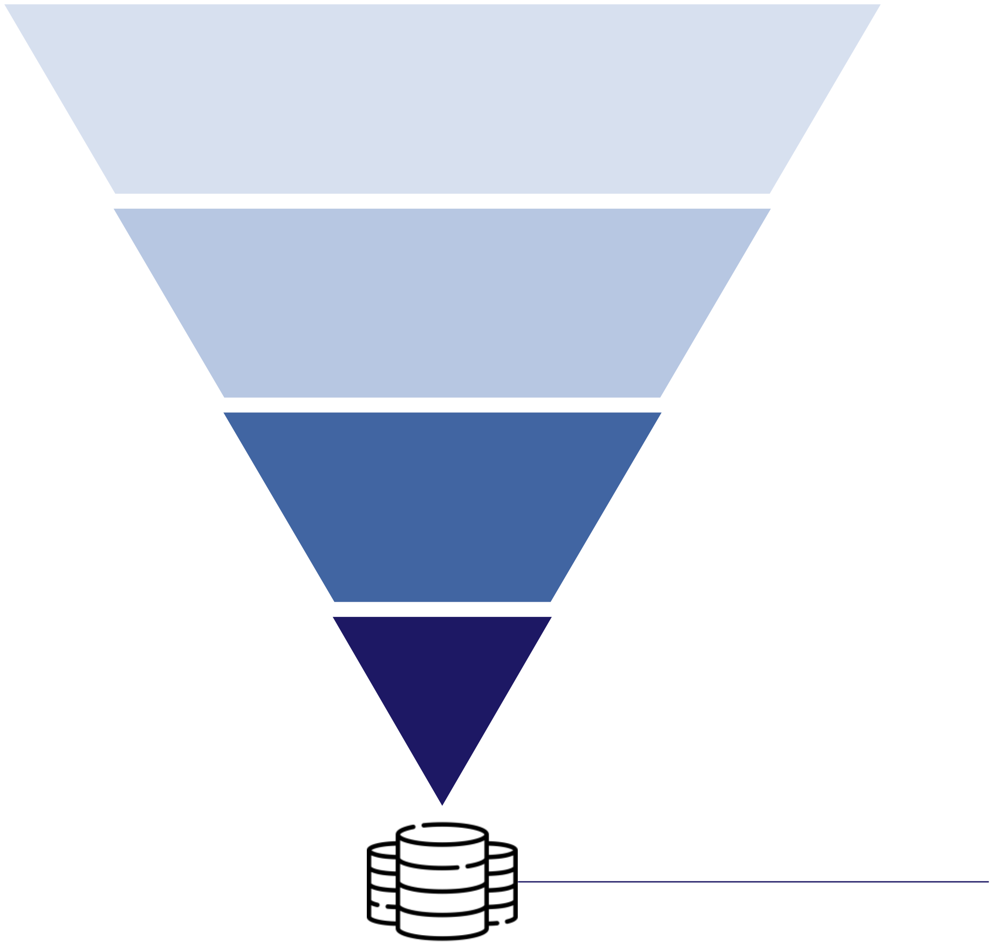
Tuning

Parameters

Validation

Adjusting model parameters to enhance accuracy and efficiency

Machine learning models can only perform in respect to their data quality



Expression of Interest

Engagement with Online tools

Formal Applicants

Final Approval

Our dataset

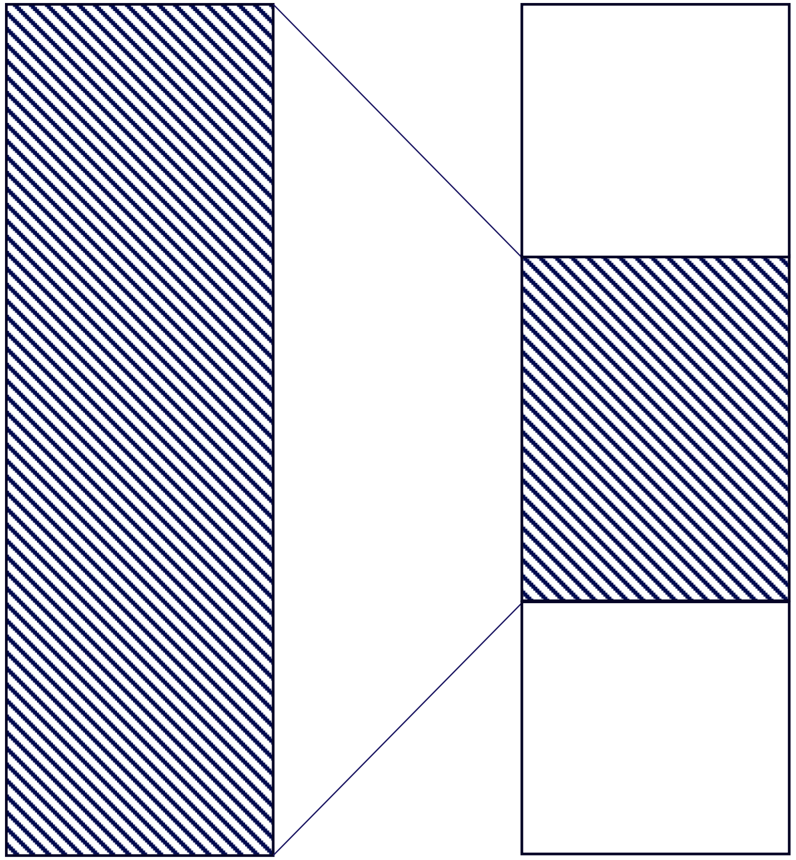
Using a dataset of only approved loans does not paint the full picture. Our models are not comparable to current legacy systems.

Results and recommendations are then inherently biased, and their real-world performance can only be stipulated.

Filtered down data is a drawback

By using a sample of the dataset, we minimise computing time

We used a sample of the dataset



The key characteristics of the dataset

0.01% | Delinquency rate

20 year | Timeframe

1M+ | Total datapoints

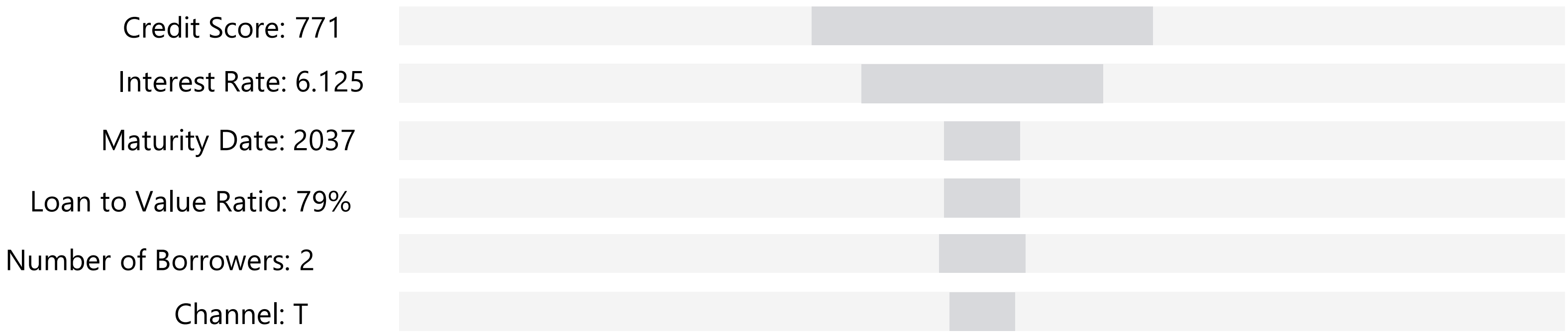
Focus 2: ML Explainability

- How do these models make their decisions?
- Why do we need to explain ML decisions?
    I. Transparency
    II. Improvement

Example of Roger

Delinquent

LIME Analysis



DiCE Analysis

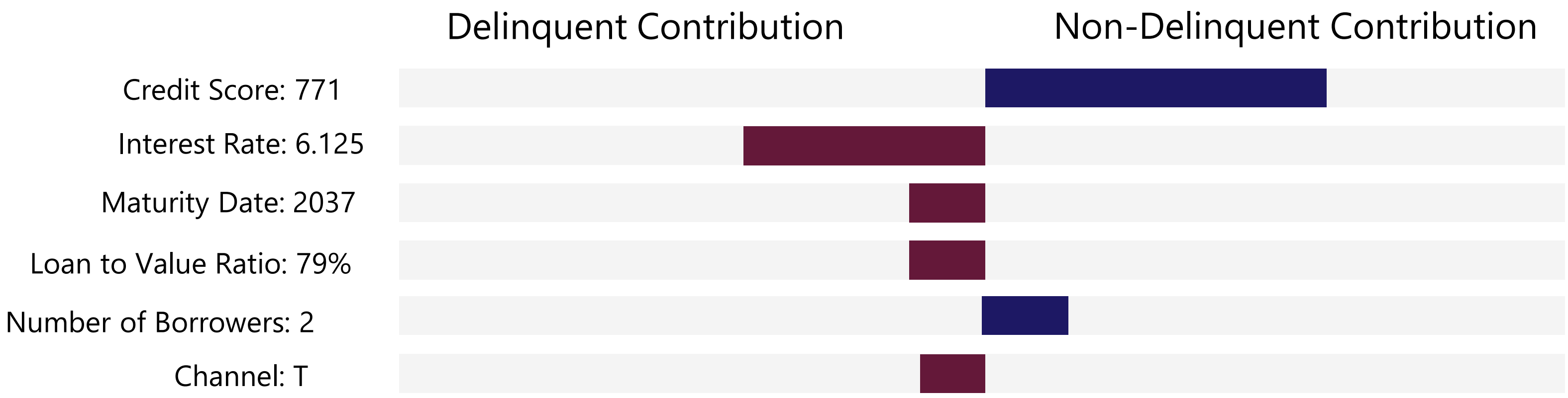
Roger	Credit Score	LTV	Interest Rate	Property State	Status
Original	771	79%	6.125%	MA	Delinquent
Example 1	-	-	4.793%	-	Non-Delinquent
Example 2	-	43%	-	-	Non-Delinquent

What can Roger do to decrease his PD?

Example of Roger

Delinquent

LIME Analysis

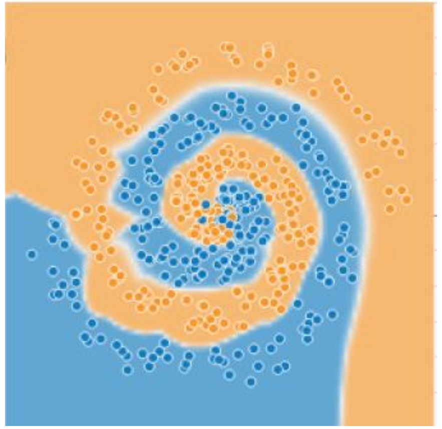


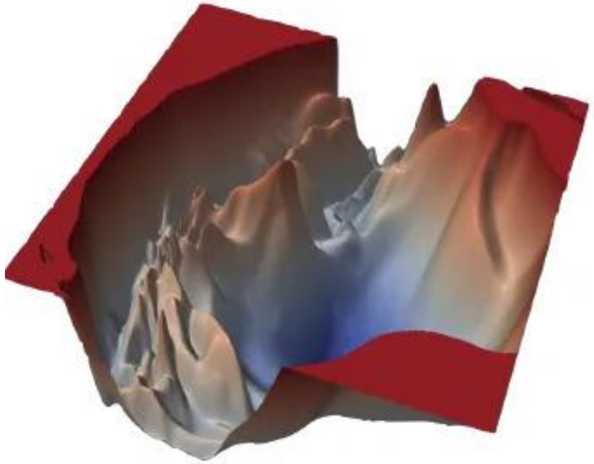
DiCE Analysis

Roger	Credit Score	LTV	Interest Rate	Property State	Status
Original	771	79%	6.125%	MA	Delinquent
Example 1	-	-	4.793%	-	Non-Delinquent
Example 2	-	43%	-	-	Non-Delinquent

What can Roger do to decrease his PD?

How does our model decide who gets a loan or not?





Decision planes

Arbitrary shape

The plane is formed through the solving of a minimization problem.

Complex

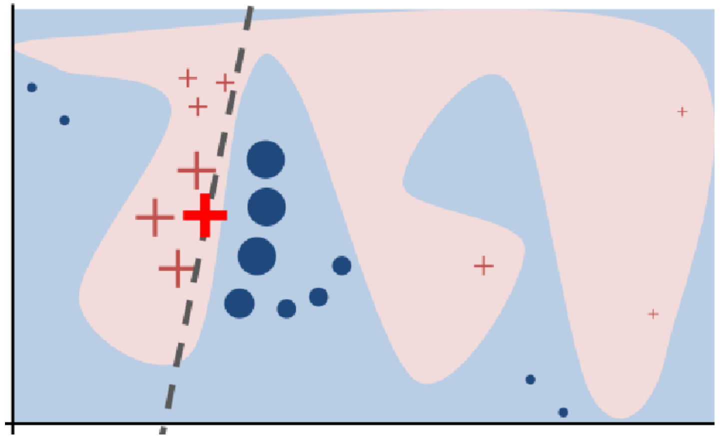
The complexity is beyond visualization.

Dimensions

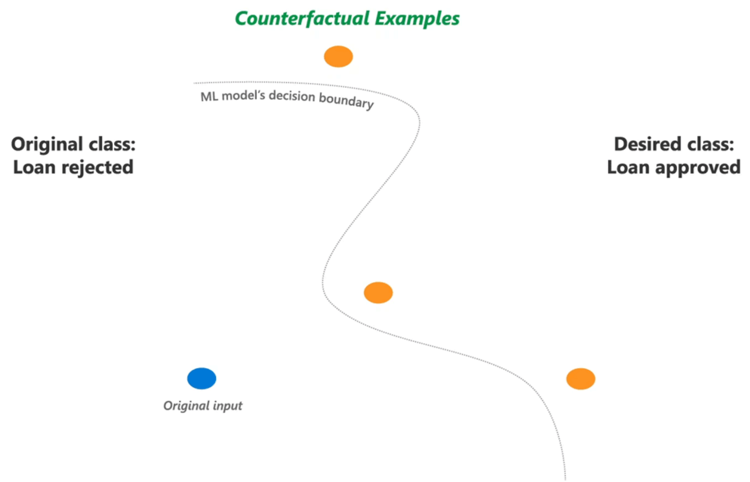
Due to the vast number of variables fed into our model, the decision plane would be multi-dimensional

Introducing LIME (Local Interpretable Model-Agnostic Explanations) and DiCE (Diverse Counterfactual Explanations)

LIME
Sampling Hypothetical Scenarios
Utilize a Simple Model
Local dynamics

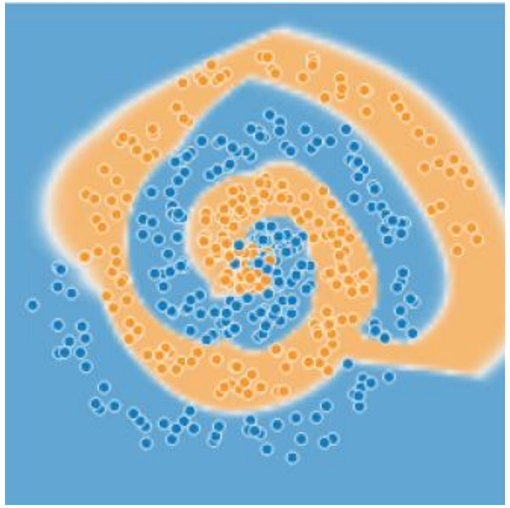


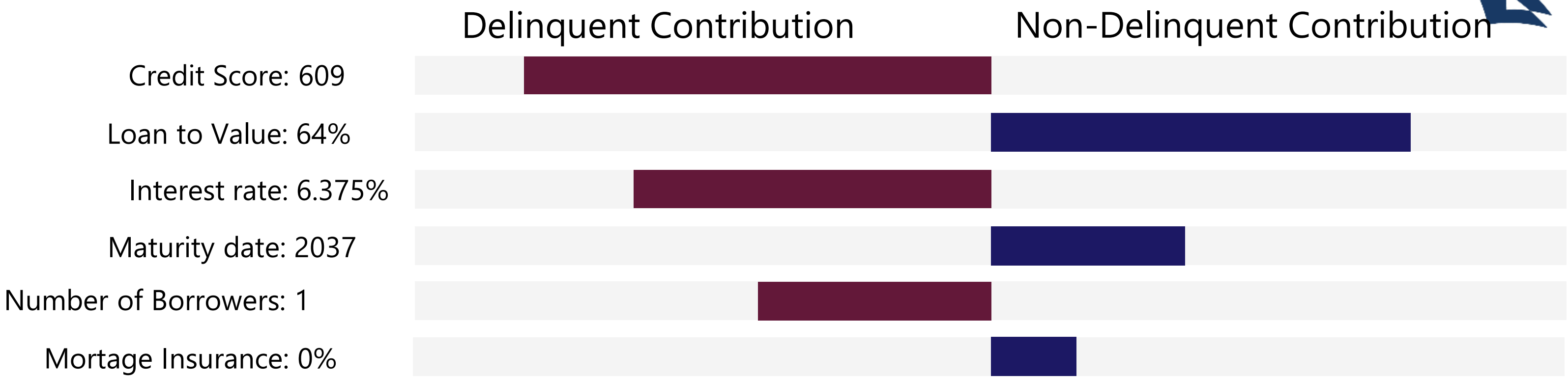
DiCE
Not only local features
What client needs to do
DiCE solves a puzzle



Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "why should I trust you?" *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. https://doi.org/10.1145/2939672.2939778

Counter intuitive example

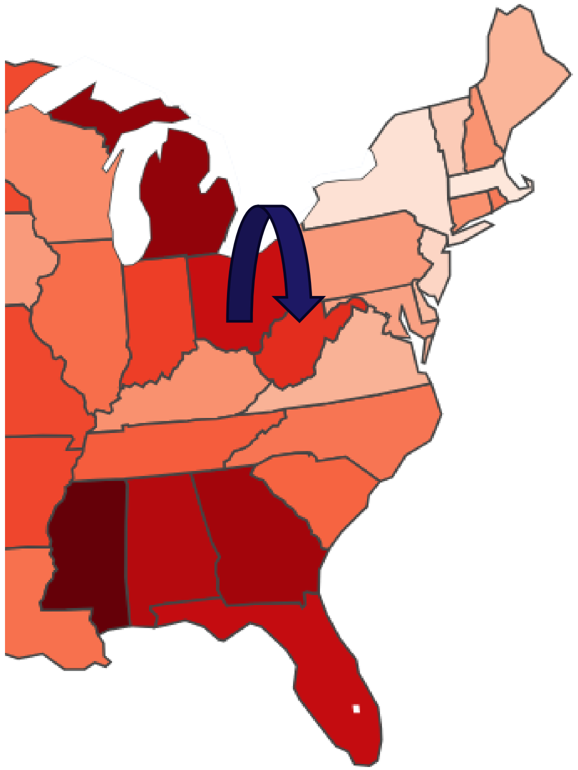




Good Accuracy ≠ Good Decisions

Client #2	Credit Score	LTV	Interest Rate	Property State	Status
Original	609	64%	6.375%	CT	Delinquent
Example 1	821	-	-	-	Non-Delinquent
Example 2	-	-	3.949%	-	Non-Delinquent
Example 3	317	-	-	-	Non-Delinquent
Example 4	-	-	12.795%	-	Non-Delinquent

Detection of Bias



Client #3	Credit Score	LTV	Interest Rate	Property State	Status
Original	723	97%	6.5%	OH	Delinquent
Example 1	-	50.9	-	-	Non-Delinquent
Example 2	-	-	-	WV	Non-Delinquent

Explainable? Yes, however….

Model Agnostic

This frameworks are easy to use and implement to any model of your choosing.

Easy to use

They may be used alongside a credit officer to give him insight on the decision-making method of the model.

Computation

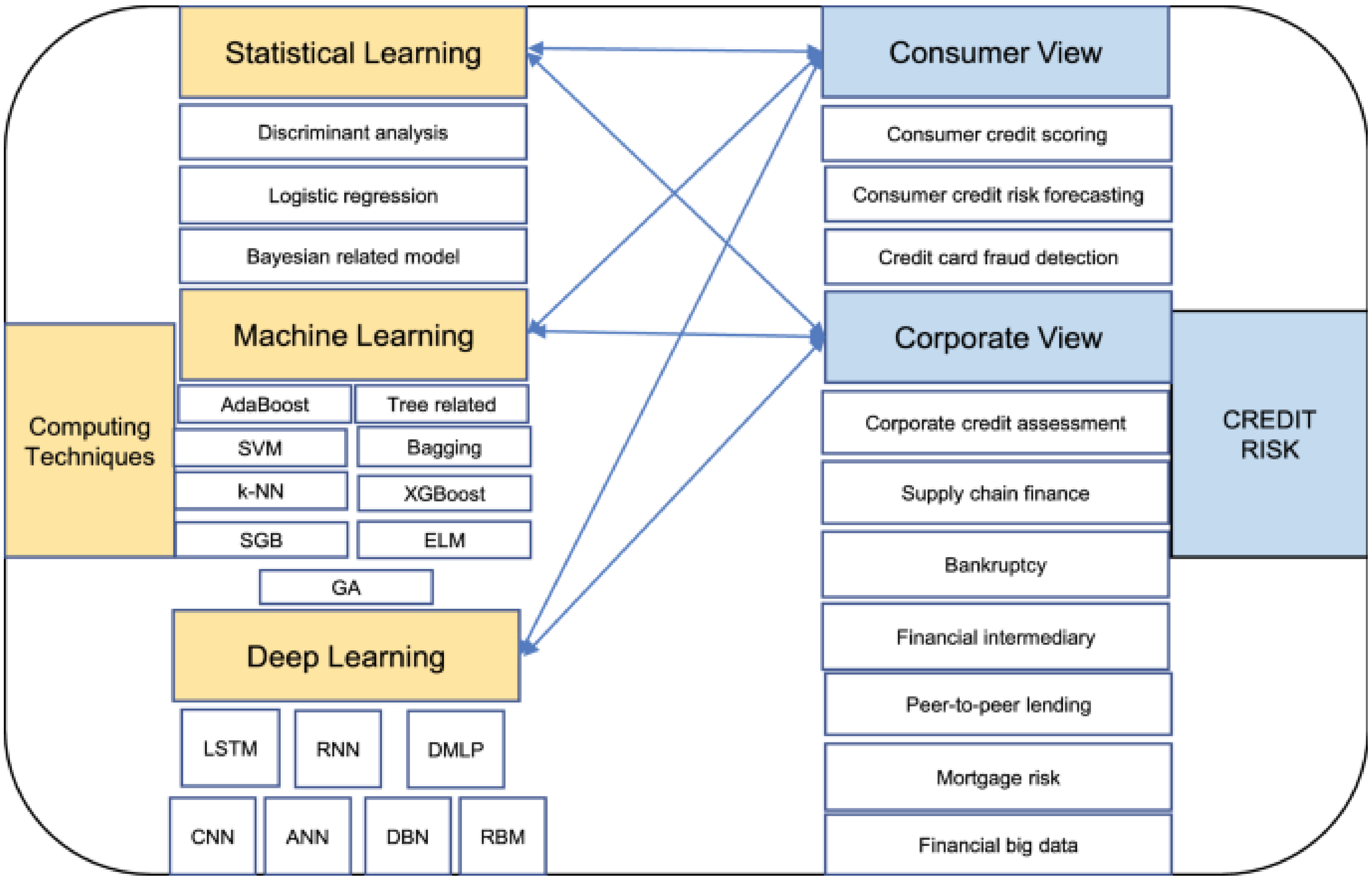
Computationally expensive, requiring precomputing and real time computing.

Human Evaluation

Suggested to be supervised by a human to ensure logical recommendations are being made.

Appendix I

Taxonomy



Statistical Learning
Discriminant analysis
Logistic regression
Bayesian related model
Machine Learning
AdaBoost
Tree related
SVM
Bagging
k-NN
XGBoost
SGB
ELM
GA
Computing Techniques
Deep Learning
LSTM
RNN
DMLP
CNN
ANN
DBN
RBM
Consumer View
Consumer credit scoring
Consumer credit risk forecasting
Credit card fraud detection
Corporate View
Corporate credit assessment
Supply chain finance
Bankruptcy
Financial intermediary
Peer-to-peer lending
Mortgage risk
Financial big data
CREDIT RISK

Optimization Loss of DiCE (Diverse Counterfactual Explanations)

$$C(\mathbf{x}) = \underset{\mathbf{c}_1,\ldots,\mathbf{c}_k}{\arg\min} \frac{1}{k}\sum_{i=1}^{k} \text{yloss}(f(\mathbf{c}_i), y) + \frac{\lambda_1}{k}\sum_{i=1}^{k} dist(\mathbf{c}_i, \mathbf{x}) - \lambda_2 \,\text{dpp\_diversity}(\mathbf{c}_1, \ldots, \mathbf{c}_k)$$

- f ➡ The original ML model (Gradient Boosted Decision Tree)
- $\mathbf{x}$ ➡ The original Input (the user's attributes)
- $\mathbf{c}_i$ ➡ The i counterfactual example
- y ➡ The opposite class label (Delinquent or not)
- $\text{yloss}(f(\mathbf{c}_i), y)$ ➡ The loss function of the opposite label and the counterfactual example
- $dist(\mathbf{c}_i, \mathbf{x})$ ➡ The distance of the counterfactual example and the original input (the user's attributes)
- dpp_diversity ➡ The Determinantal Point Processes of the kernelized counterfactual exampes (how diverse they are)

Mothilal, R. K., Sharma, A., & Tan, C. (2020). Explaining machine learning classifiers through diverse counterfactual explanations. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency.* https://doi.org/10.1145/3351095.3372850

Optimization Loss of DiCE (Diverse Counterfactual Explanations)

$$C(\boldsymbol{x}) = \arg\min_{\boldsymbol{c}_1,\ldots,\boldsymbol{c}_k} \frac{1}{k}\sum_{i=1}^{k} \mathrm{yloss}(f(\boldsymbol{c}_i), y) + \frac{\lambda_1}{k}\sum_{i=1}^{k} dist(\boldsymbol{c}_i, \boldsymbol{x}) - \lambda_2\, \mathrm{dpp\_diversity}(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_k)$$

f ➡ The original ML model (Gradient Boosted Decision Tree)

$\boldsymbol{x}$ ➡ Vector of size the number of the model's features

$\boldsymbol{c}_i$ ➡ Vector of size similar to the input, the number of the model's features

y ➡ Scalar value of 0 or 1

$$\mathrm{yloss}(f(\boldsymbol{c}_i), y) \Rightarrow hinge\_yloss = max(0, 1 - z * logit(f(\boldsymbol{c})))$$

$$dist(\boldsymbol{c}_i, \boldsymbol{x}) \Rightarrow \mathrm{dist\_cont}(\boldsymbol{c}, \boldsymbol{x}) = \frac{1}{d_{cont}} \sum_{p=1}^{d_{cont}} \frac{|\boldsymbol{c}^p - \boldsymbol{x}^p|}{MAD_p} \quad \mathrm{dist\_cat}(\boldsymbol{c}, \boldsymbol{x}) = \frac{1}{d_{cat}} \sum_{p=1}^{d_{cat}} I(\boldsymbol{c}^p \neq \boldsymbol{x}^p)$$

Mothilal, R. K., Sharma, A., & Tan, C. (2020). Explaining machine learning classifiers through diverse counterfactual explanations. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency.* https://doi.org/10.1145/3351095.3372850

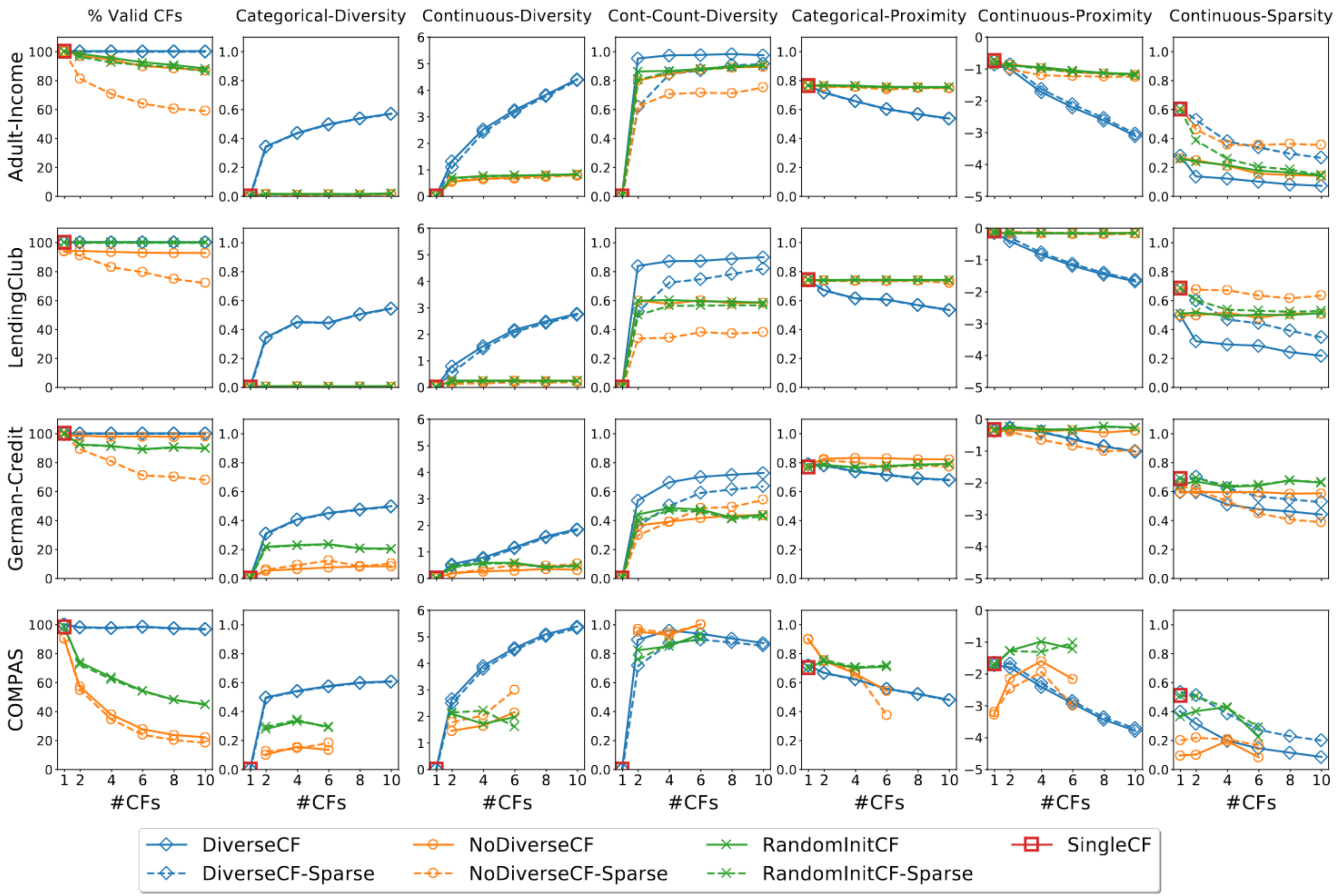
Optimization tricks used by DiCE (Diverse Counterfactual Explanations)

$$C(\boldsymbol{x}) = \underset{\boldsymbol{c}_1,\ldots,\boldsymbol{c}_k}{\arg\min} \frac{1}{k}\sum_{i=1}^{k} \text{yloss}(f(\boldsymbol{c}_i), y) + \frac{\lambda_1}{k}\sum_{i=1}^{k} dist(\boldsymbol{c}_i, \boldsymbol{x}) - \lambda_2 \,\text{dpp\_diversity}(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_k)$$

- Sparsity
  - They restore the value of continuous features back to their values in x greedily if less than a chosen threshold.
- Hyperparameters
  - The hyperparameter λ1 balances the counterfactuals distance from the origin and the λ2 balances their diversity. Both were derived by using Grid Search
- MAD
  - They use Median Absolute Deviation for more robust distance measures

Mothilal, R. K., Sharma, A., & Tan, C. (2020). Explaining machine learning classifiers through diverse counterfactual explanations. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency.* https://doi.org/10.1145/3351095.3372850

Results of DiCE (Diverse Counterfactual Explanations)



Mothilal, R. K., Sharma, A., & Tan, C. (2020). Explaining machine learning classifiers through diverse counterfactual explanations.
Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency. https://doi.org/10.1145/3351095.3372850

Optimization Loss of LIME (Local Interpretable Model-agnostic Explanations)

$$\xi(x) = \underset{g \in G}{\operatorname{argmin}} \ \mathcal{L}(f, g, \pi_x) + \Omega(g)$$

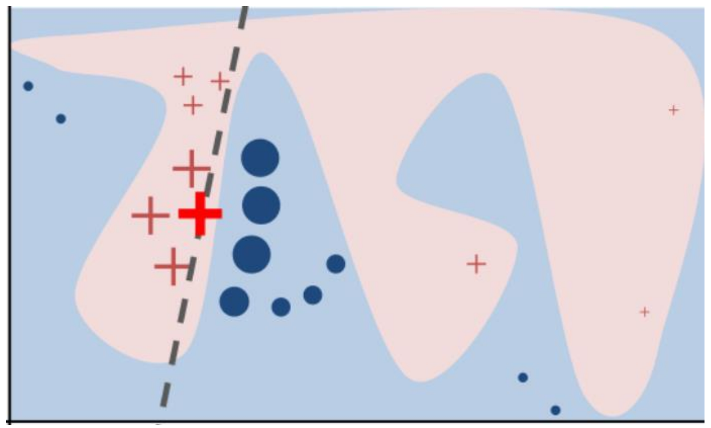
- f ➡ The original ML model (Gradient Boosted Decision Tree)
- g ➡ Simpler Model (Linear model to explain locally the more complex one)
- π_x ➡ Sampled points near input x
- $\Omega(g)$ ➡ Regularization factor
- $\mathcal{L}(f, g, \pi_x)$ ➡ The loss function how well the simple model approximates the complex one

$$\mathcal{L}(f, g, \pi_x) = \sum_{z, z' \in \mathcal{Z}} \pi_x(z) \left(f(z) - g(z') \right)^2$$

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "why should I trust you?" *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. https://doi.org/10.1145/2939672.2939778

Optimization loss of LIME (Local Interpretable Model-agnostic Explanations)

$$\xi(x) = \underset{g \in G}{\operatorname{argmin}} \; \mathcal{L}(f, g, \pi_x) + \Omega(g)$$

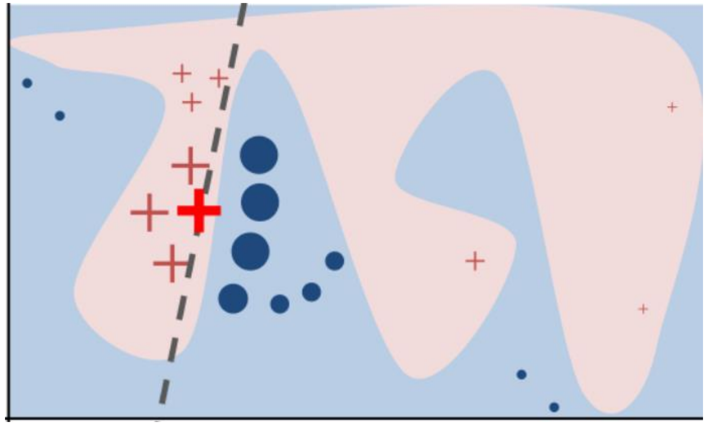


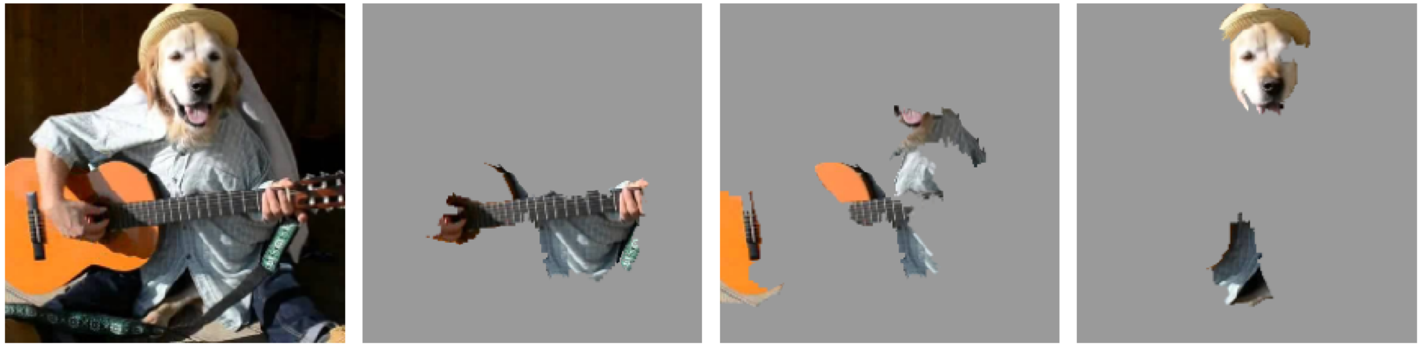
- RBF kernel is used to measure the distance of the sampled points $\pi_x(z) = exp(-D(x,z)^2/\sigma^2)$
- 5000 points are sampled by default. We sampled 50000
- Model agnostic

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "why should I trust you?" *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. https://doi.org/10.1145/2939672.2939778

Results of LIME in images (Local Interpretable Model-agnostic Explanations)

(a) Husky classified as wolf (b) Explanation





(a) Original Image (b) Explaining *Electric guitar* (c) Explaining *Acoustic guitar* (d) Explaining *Labrador*

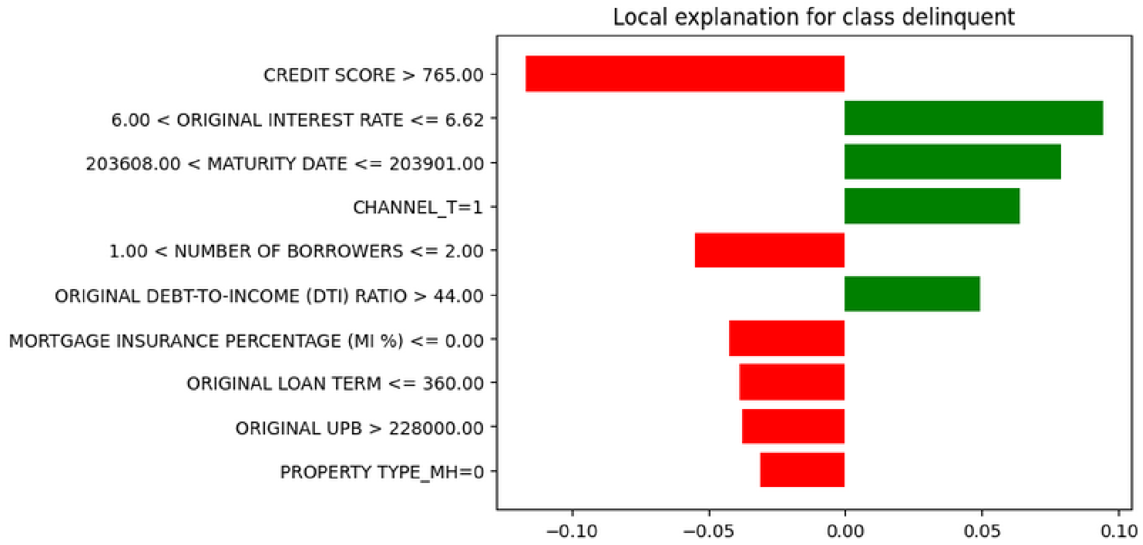
Figure 4: Explaining an image classification prediction made by Google's Inception neural network. The top 3 classes predicted are "Electric Guitar" ($p = 0.32$), "Acoustic guitar" ($p = 0.24$) and "Labrador" ($p = 0.21$)

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "why should I trust you?" *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* https://doi.org/10.1145/2939672.2939778

Our counterfactual examples on the dataset by using the Gradient Boosted Tree

Example 1

LIME



DiCE

	CREDIT SCORE	MATURITY DATE	MORTGAGE INSURANCE PERCENTAGE (MI %)	NUMBER OF UNITS	ORIGINAL COMBINED LOAN-TO-VALUE (CLTV)	ORIGINAL DEBT-TO-INCOME (DTI) RATIO	ORIGINAL UPB	ORIGINAL LOAN-TO-VALUE (LTV)	ORIGINAL INTEREST RATE	ORIGINAL LOAN TERM	NUMBER OF BORROWERS	LATITUDE	LONGITUDE	FIRST TIME HOMEBUYER FLAG	OCCUPANCY STATUS	CHANNEL	PROPERTY STATE	PROPERTY TYPE	LOAN PURPOSE	CURRENT LOAN DELINQUENCY STATUS
0	771.0	203701.0	0.0	1.0	79.0	52.0	330000.0	79.0	6.125	360.0	2.0	41.70401	-70.228828	N	P	T	MA	SF	C	1.0

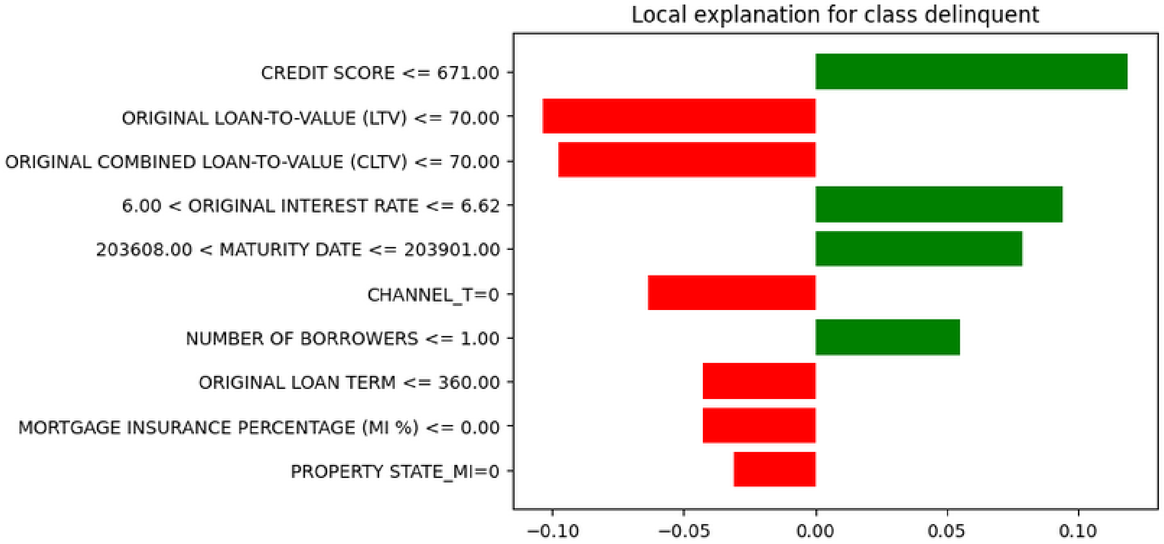
Diverse Counterfactual set (new outcome: 0.0)

	CREDIT SCORE	MATURITY DATE	MORTGAGE INSURANCE PERCENTAGE (MI %)	NUMBER OF UNITS	ORIGINAL COMBINED LOAN-TO-VALUE (CLTV)	ORIGINAL DEBT-TO-INCOME (DTI) RATIO	ORIGINAL UPB	ORIGINAL LOAN-TO-VALUE (LTV)	ORIGINAL INTEREST RATE	ORIGINAL LOAN TERM	NUMBER OF BORROWERS	LATITUDE	LONGITUDE	FIRST TIME HOMEBUYER FLAG	OCCUPANCY STATUS	CHANNEL	PROPERTY STATE	PROPERTY TYPE	LOAN PURPOSE	CURRENT LOAN DELINQUENCY STATUS
0	-	-	-	-	-	-	-	43.0	-	-	-	-	-	-	-	-	-	-	-	0.0
1	-	-	-	-	-	-	-	-	4.793	-	-	-	-	-	-	-	-	-	-	0.0
2	-	-	-	-	10.2	-	-	24.0	-	-	-	-	-	-	-	-	-	-	-	0.0
3	-	-	-	-	48.7	-	-	120.6	-	-	-	-	-	-	-	-	-	-	-	0.0
4	-	-	-	-	-	18.1	-	140.1	-	-	-	-	-	-	-	-	-	-	-	0.0
5	-	-	-	-	-	-	-	-	1.251	-	-	-	-	-	-	-	-	-	-	0.0
6	-	-	-	-	-	-	-	-	3.864	-	-	-	-	-	-	-	-	-	-	0.0
7	-	-	-	-	38.1	-	-	-	2.694	-	-	-	-	-	-	-	-	-	-	0.0
8	-	-	-	-	-	-	-	24.0	-	-	-	-	-	-	-	-	-	-	-	0.0
9	-	-	-	-	-	-	-	42.3	1.251	-	-	-	-	-	-	-	-	-	-	0.0

Our counterfactual examples on the dataset by using the Gradient Boosted Tree

Example 2

LIME



	CREDIT SCORE	MATURITY DATE	MORTGAGE INSURANCE PERCENTAGE (MI %)	NUMBER OF UNITS	ORIGINAL COMBINED LOAN-TO-VALUE (CLTV)	ORIGINAL DEBT-TO-INCOME (DTI) RATIO	ORIGINAL UPB	ORIGINAL LOAN-TO-VALUE (LTV)	ORIGINAL INTEREST RATE	ORIGINAL LOAN TERM	NUMBER OF BORROWERS	LATITUDE	LONGITUDE	FIRST TIME HOMEBUYER FLAG	OCCUPANCY STATUS	CHANNEL	PROPERTY STATE	PROPERTY TYPE	LOAN PURPOSE	CURRENT LOAN DELINQUENCY STATUS
0	609.0	203708.0	0.0	1.0	64.0	30.0	127000.0	64.0	6.375	360.0	1.0	41.306881	-72.466522	N	P	R	CT	SF	C	1.0

DiCE

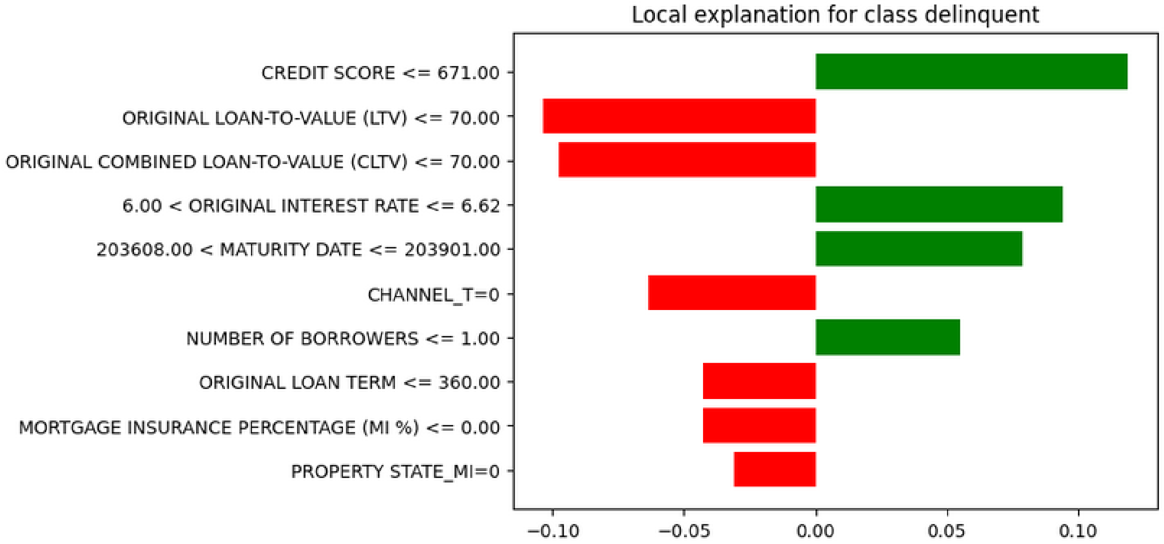
Diverse Counterfactual set (new outcome: 0.0)

	CREDIT SCORE	MATURITY DATE	MORTGAGE INSURANCE PERCENTAGE (MI %)	NUMBER OF UNITS	ORIGINAL COMBINED LOAN-TO-VALUE (CLTV)	ORIGINAL DEBT-TO-INCOME (DTI) RATIO	ORIGINAL UPB	ORIGINAL LOAN-TO-VALUE (LTV)	ORIGINAL INTEREST RATE	ORIGINAL LOAN TERM	NUMBER OF BORROWERS	LATITUDE	LONGITUDE	FIRST TIME HOMEBUYER FLAG	OCCUPANCY STATUS	CHANNEL	PROPERTY STATE	PROPERTY TYPE	LOAN PURPOSE	CURRENT LOAN DELINQUENCY STATUS
0	-	-	-	-	-	-	-	35.7	4.974	-	-	-	-	-	-	-	-	-	-	0.0
1	-	-	-	-	-	-	-	-	3.949	-	-	-	-	-	-	-	-	-	-	0.0
2	319.8	-	-	-	-	-	-	-	0.6	-	-	-	-	-	-	-	-	-	-	0.0
3	-	-	-	-	-	-	-	22.3	3.088	-	-	-	-	-	-	-	-	-	-	0.0
4	-	-	-	-	24.7	-	-	24.7	-	-	-	-	-	-	-	-	-	-	-	0.0
5	821.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0
6	-	-	-	-	-	4.7	-	-	1.111	-	-	-	-	-	-	-	-	-	-	0.0
7	-	-	-	-	-	-	-	-	1.144	-	-	-	-	-	-	-	-	-	-	0.0
8	-	-	-	-	-	-	-	-	0.477	-	-	-	-	-	-	-	-	-	-	0.0
9	-	-	-	-	-	13.9	-	-	3.06	-	-	-	-	-	-	-	-	-	-	0.0

Our counterfactual examples on the dataset by using the Gradient Boosted Tree

Bad Example 1

LIME



	CREDIT SCORE	MATURITY DATE	MORTGAGE INSURANCE PERCENTAGE (MI %)	NUMBER OF UNITS	ORIGINAL COMBINED LOAN-TO-VALUE (CLTV)	ORIGINAL DEBT-TO-INCOME (DTI) RATIO	ORIGINAL UPB	ORIGINAL LOAN-TO-VALUE (LTV)	ORIGINAL INTEREST RATE	ORIGINAL LOAN TERM	NUMBER OF BORROWERS	LATITUDE	LONGITUDE	FIRST TIME HOMEBUYER FLAG	OCCUPANCY STATUS	CHANNEL	PROPERTY STATE	PROPERTY TYPE	LOAN PURPOSE	CURRENT LOAN DELINQUENCY STATUS
0	609.0	203708.0	0.0	1.0	64.0	30.0	127000.0	64.0	6.375	360.0	1.0	41.306881	-72.466522	N	P	R	CT	SF	C	1.0

DiCE

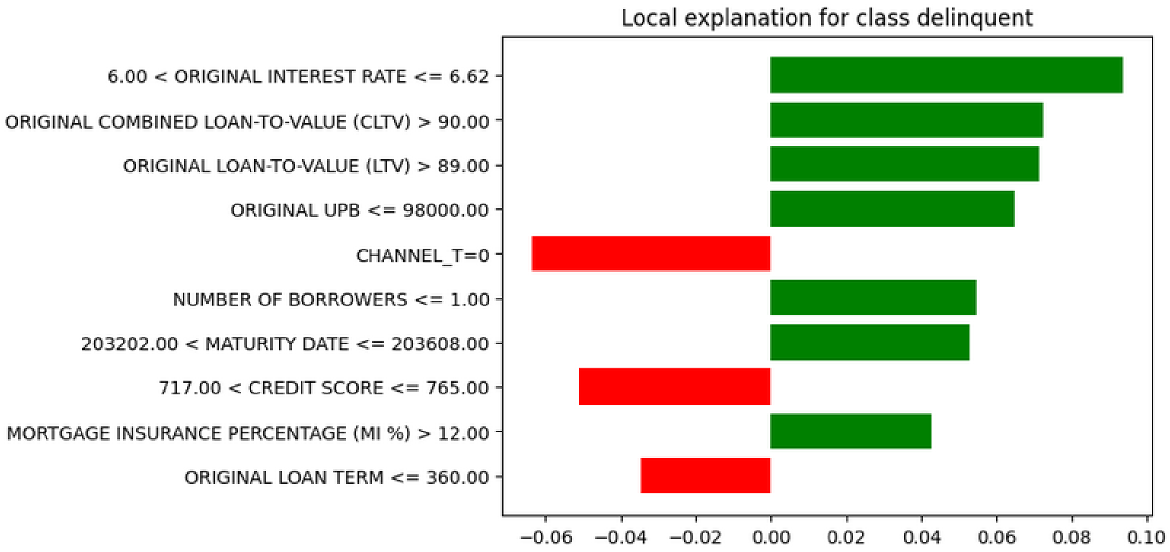
Diverse Counterfactual set (new outcome: 0.0)

	CREDIT SCORE	MATURITY DATE	MORTGAGE INSURANCE PERCENTAGE (MI %)	NUMBER OF UNITS	ORIGINAL COMBINED LOAN-TO-VALUE (CLTV)	ORIGINAL DEBT-TO-INCOME (DTI) RATIO	ORIGINAL UPB	ORIGINAL LOAN-TO-VALUE (LTV)	ORIGINAL INTEREST RATE	ORIGINAL LOAN TERM	NUMBER OF BORROWERS	LATITUDE	LONGITUDE	FIRST TIME HOMEBUYER FLAG	OCCUPANCY STATUS	CHANNEL	PROPERTY STATE	PROPERTY TYPE	LOAN PURPOSE	CURRENT LOAN DELINQUENCY STATUS
0	-	-	-	-	-	-	-	-	12.795	-	-	-	-	-	-	-	-	-	-	0.0
1	828.7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0
2	-	-	-	-	-	-	-	-	4.175	-	-	-	-	-	-	-	-	-	-	0.0
3	-	-	-	-	-	-	-	-	1.734	-	-	-	-	-	-	-	-	-	-	0.0
4	-	-	-	-	-	-	-	-	11.307	-	-	-	-	-	-	-	-	-	-	0.0
5	-	-	-	-	-	-	-	-	7.372	-	-	-	-	-	-	-	-	-	-	0.0
6	-	-	-	-	-	-	-	-	4.756	-	-	-	-	-	-	-	-	-	-	0.0
7	835.3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0
8	317.2	-	-	-	-	-	-	-	0.533	-	-	-	-	-	-	-	-	-	-	0.0
9	-	-	-	-	-	-	-	-	5.048	-	-	-	-	-	-	-	-	-	-	0.0

Our counterfactual examples on the dataset by using the Gradient Boosted Tree

Bad Biased Example 2

LIME



	CREDIT SCORE	MATURITY DATE	MORTGAGE INSURANCE PERCENTAGE (MI %)	NUMBER OF UNITS	ORIGINAL COMBINED LOAN-TO-VALUE (CLTV)	ORIGINAL DEBT-TO-INCOME (DTI) RATIO	ORIGINAL UPB	ORIGINAL LOAN-TO-VALUE (LTV)	ORIGINAL INTEREST RATE	ORIGINAL LOAN TERM	NUMBER OF BORROWERS	LATITUDE	LONGITUDE	FIRST TIME HOMEBUYER FLAG	OCCUPANCY STATUS	CHANNEL	PROPERTY STATE	PROPERTY TYPE	LOAN PURPOSE	CURRENT LOAN DELINQUENCY STATUS
0	723.0	203310.0	35.0	1.0	97.0	39.0	74000.0	97.0	6.5	360.0	1.0	38.936371	-87.269592	Y	P	R	OH	SF	P	1.0

DiCE

Diverse Counterfactual set (new outcome: 0.0)

	CREDIT SCORE	MATURITY DATE	MORTGAGE INSURANCE PERCENTAGE (MI %)	NUMBER OF UNITS	ORIGINAL COMBINED LOAN-TO-VALUE (CLTV)	ORIGINAL DEBT-TO-INCOME (DTI) RATIO	ORIGINAL UPB	ORIGINAL LOAN-TO-VALUE (LTV)	ORIGINAL INTEREST RATE	ORIGINAL LOAN TERM	NUMBER OF BORROWERS	LATITUDE	LONGITUDE	FIRST TIME HOMEBUYER FLAG	OCCUPANCY STATUS	CHANNEL	PROPERTY STATE	PROPERTY TYPE	LOAN PURPOSE	CURRENT LOAN DELINQUENCY STATUS
0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	WV	-	-	0.0
1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0
2	-	-	-	-	50.9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0
3	-	-	-	-	-	-	-	-	0.5	-	-	24.48083	-	-	-	-	-	-	-	0.0
4	-	-	8.1	-	-	-	-	-	2.341	-	-	-	-	-	-	-	-	-	-	0.0
5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.0
6	-	-	-	-	-	-	-	-	3.973	-	-	-	-	-	-	-	-	-	-	0.0
7	-	-	-	-	-	-	-	17.2	-	-	2.5	-	-	-	-	-	-	-	-	0.0
8	-	-	-	-	-	-	-	-	3.186	-	-	-	-	-	-	-	-	-	-	0.0
9	-	-	-	-	-	-	-	39.0	-	-	4.5	-	-	-	-	-	-	-	-	0.0