UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II

Artificial Intelligence and Intelligent Systems
cini National Lab

PICUS lab
PATTERN ANALYSIS AND INTELLIGENT COMPUTATION FOR MULTIMEDIA SYSTEMS

A benchmark of credit score prediction using Machine Learning

Vincenzo Moscato and **Giancarlo Sperlì**

Ital-IA 2023: 3rd National Conference on Artificial Intelligence May 29--31, 2023, Pisa, Italy
AI per la Finanza ed il Commercio

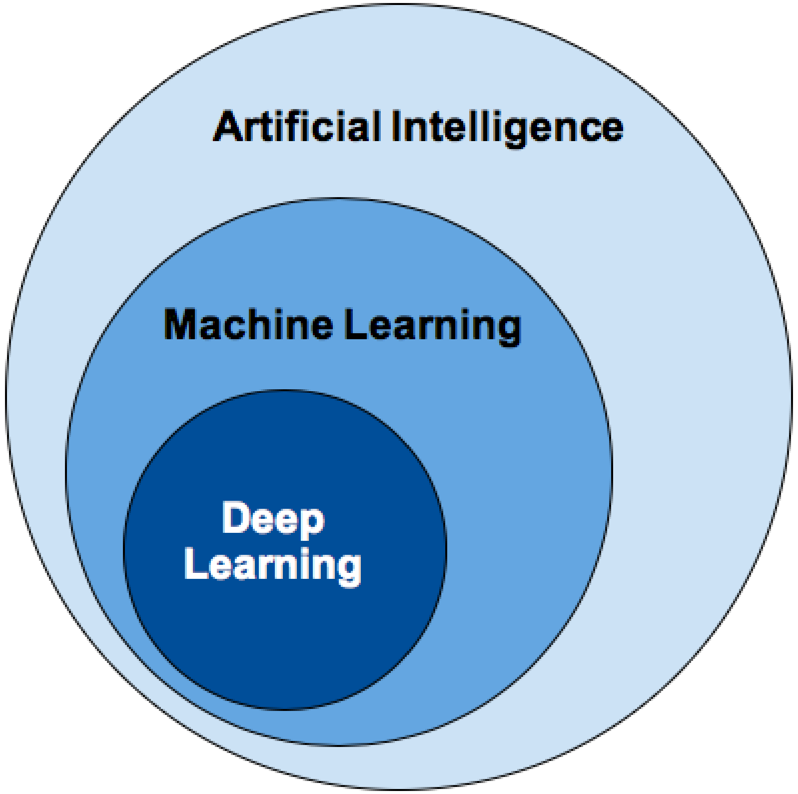
Artificial Intelligence in Finance

- Several financial services have benefited from the introduction of Artificial Intelligence (AI)-based models by defining a new generation of financial technology (FinTech)-based systems.

- AI in Finance can be used to:
  - ✓ Support payment processes;
  - ✓ Analyze people history for credit scoring;
  - ✓ Support risk and regulatory management.

- The most common application of traditional AI is credit scoring:
  - ✓ predicting whether debts are repaid or not (binary problem)

- Challenges:
  - ✓ Big Data Analysis:
    - ➢ One of the main challenge is the large amount of data produced by digital financial services.
  - ✓ Risk management:
    - ➢ Risks can be classified into three categories, namely credit, market, and operational risks.



Social Lending Platform

- In the last years, traditional credit risk services have been disrupted by the arise of Social Lending Platforms.
- Social Lending Platforms enable communications among lenders and borrowers without any transaction costs, that are typically for traditional financial institute.
- Lenders are exposed to risks when investing in Social Lending Platform, particularly in the form of credit risk, which is assessed through the process of credit scoring. This risk arises primarily from the possibility that borrowers may be unable to repay their loans.
- Credit risks account for approximately 60% of banks' risks, which is mainly due to the arise of Social Lending Platforms.
- Different statistical approaches have been proposed although they do not properly cover non-linear effects among different variables.

Social Lending Platform - Challenges

- Social Lending Platforms facilitate the fundraising process for borrowers by allowing lenders of all sizes to participate.

- Social lending platforms pose unique challenges with respect to traditional methods, dealing with:
  - ✓ <u>high-dimensionality,</u>
  - ✓ <u>sparsity,</u>
  - ✓ <u>imbalance data</u>

- The risk of defaults in P2P lending platforms is generally higher than in traditional methods due to the issues of lenders in accurately assessing borrowers' risk levels;

- The primary challenge concerns how it is possible to evaluate creditworthiness of loan applicants, since borrowers often lack a sufficient credit history, and simply adding more features may not necessarily improve the accuracy of the assessment.

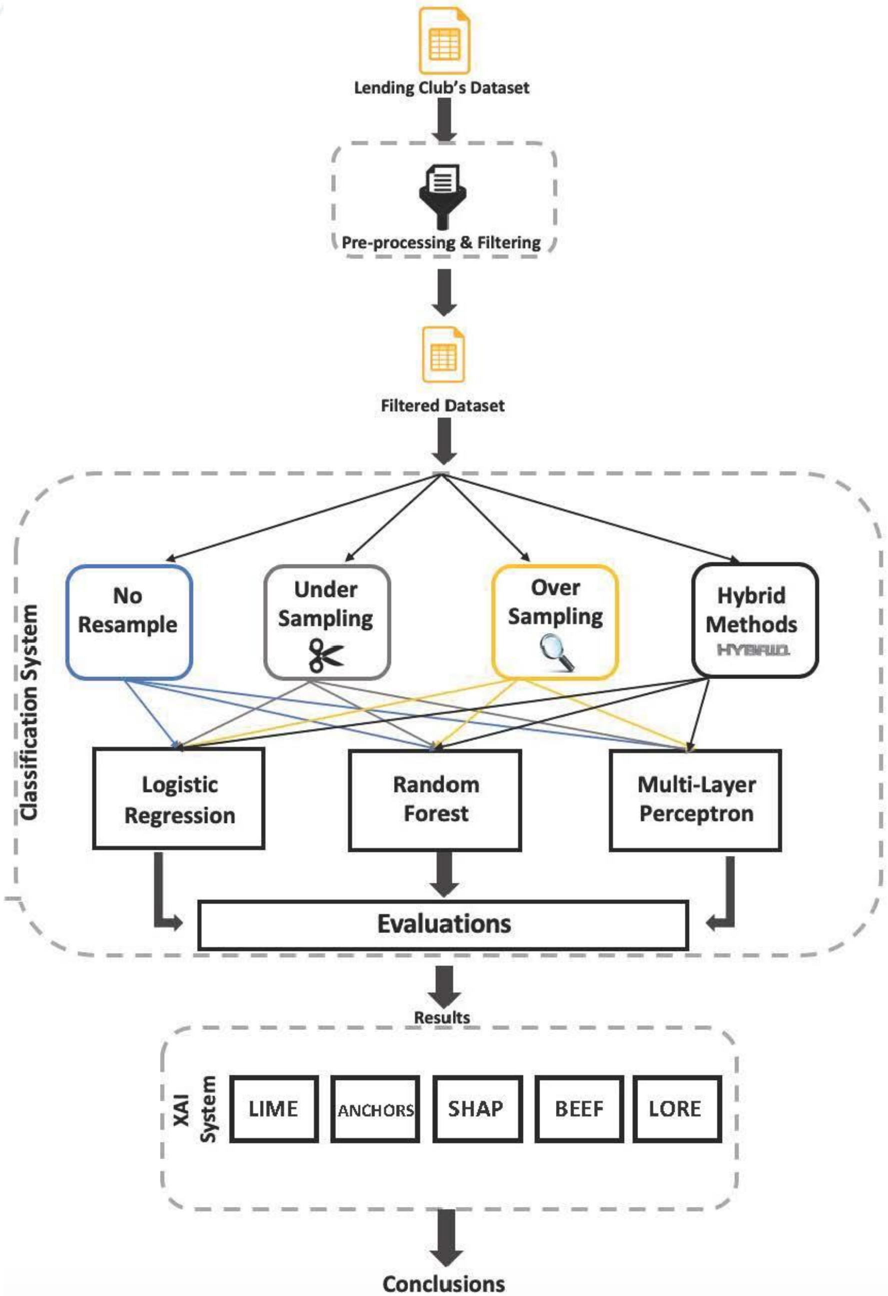
Social Lending Platform - Proposal

- One of the main relevant financial services is the credit risk assessment, whose aim is to support financial institutes in defining their policies and strategies.

- In Social Lending Platform, lenders can earn higher returns than what is typically offered through banks' savings and investment products, while borrowers can access funds at lower interest rates.

- To deal with the previous defined issues, we:
  - ✓ Designed a benchmark of machine learning models for credit scoring prediction;
  - ✓ Investigated eXplainable Artificial Intelligence (XAI) tools for explaining the prediction of the analyzed machine learning models.
  - ✓ Evaluated both benchmark and XAI tools on a real-world Social Lending Platform, composed by 877,956 samples.

Methodology



Methodology - Modules

- The proposed benchmark is designed to deal with the credit risk prediction task with the aim to support investors in evaluating potential borrowers on Social Lending Platforms. It is composed by three main modules.

- The ingestion module is responsible for crawling data from Social Lending Platforms, also performing data cleaning and feature selection operations on the basis of the chosen classifier.

- The second component is responsible for credit prediction for a given user, which is impacted by the imbalance problem, typical issue in Social Lending platforms.
  - ✓ For the classification stage, three of most efficacy models in credit score prediction have been selected, also considering different sampling strategies.

- The third module deals with comparing different XAI techniques to explain the results obtained with the aim of explaining prediction outcome to highlight how decisions are made.

Methodology - Evaluation

- The proposed methodology has been evaluated on a real-world Social Lending Platform.
- We perform a 10-fold cross-validation, in which we split the dataset according to 75:25 ratio for each fold, computing mean and standard deviation for each classifier during the training process.
- Different measures have been used (Precision, FP-Rate, Area Under Curve (AUC), Accuracy (ACC), Sensitivity (TPR), Specificity (TNR), and G-mean).
- Baselines:
  - ✓ Logistic Regression
  - ✓ Random Forest
  - ✓ Multi-Layer Perceptron
- XAI tools:
  - ✓ LIME
  - ✓ ANCHORS
  - ✓ SHAP
  - ✓ BEEF
  - ✓ LORE

Loan Status	**Samples number**
Current	395.901
Fully Paid	354,994
Charged Off	107,384
Late (31-120 days)	12,550
In-grace period	4,703
Late (16-30 days)	2,393
Default	31
Total	**877,956**

Methodology - Prediction Results

Classifier	AUC	TPR	TNR	FP-Rate	G-Mean	ACC
RF - RUS	**0.717**	**0.630**	**0.680**	**0.320**	**0.6560**	**0.640**
LR - ROS	0.710	0.659	0.642	0.360	0.6503	0.650
LR - SmoteToken	0.710	0.660	0.640	0.360	0.6500	0.656
Logistic Regression	0.685	0.983	0.069	0.960	0.2600	0.770
Random Forest	0.720	0.983	0.084	0.920	0.2870	0.773
MLP	0.704	0.990	0.040	0.945	0.2060	0.771

Our best Classification results.

Classifier	AUC	TPR	TNR	FP-Rate	G-Mean	Accuracy
RF - RUS	**0.6960**	**0.717**	**0.582**	**0.420**	**0.65**	**0.6920**
Linear Discrimination Analysis - SMOTE	0.7000	0.630	0.650	0.350	0.643	0.6400
LR - SmoteToken	0.7000	0.638	0.648	0.352	0.643	0.6400
Logistic Regression	0.7030	0.988	0.048	0.950	0.218	0.8173
Random Forest	0.6960	0.996	0.015	0.980	0.12	0.8176

Best result in (Namvar et al. (2018))

	Method	AUC	TPR	TNR	G-Mean	Accuracy
	Song et al. (2020)	0.6697	0.4607	0.7678	0.6009	0.7231
Over-Sampling	GBDT	0.6207	0.6168	0.6246	0.6207	0.6235
	Random Forest	0.5795	0.3107	0.8423	0.5134	0.7701
	AdaBoost	0.5224	0.1925	0.8523	0.4050	0.7562
	Decision Tree	0.5231	0.1934	0.8527	0.4060	0.7568
	Logistic Regression	0.5600	0.5558	0.5642	0.5597	0.5630
	Multilayer Perceptron	0.4892	0.1572	0.8211	0.3593	0.7245
Under-Sampling	GBDT	0.6140	0.6292	0.5989	0.6138	0.6033
	Random Forest	0.6207	0.6623	0.5791	0.6193	0.5912
	AdaBoost	0.5408	0.5577	0.5238	0.5404	0.5288
	Decision Tree	0.5421	0.5558	0.5283	0.5418	0.5323
	Logistic Regression	0.5615	0.5437	0.5794	0.5609	0.5742
	Multilayer Perceptron	0.4892	0.1572	0.8211	0.3593	0.7245

Result in (Song et al. (2020))

Methodology - XAI Results

- The last evaluation has concerned the performance comparison of several XAI tools in terms of Precision measure -- what fraction of the predictions were correct -- on the our three best classifiers' combinations.

- To simulate trust on an individual prediction, we randomly chose a group of possible features (25% of the total) that must be consider “untrustworthy”, assuming that a user, that can recognize them, does not want to trust on these features.

	Random -Forest Random Under-Sampling (Precision Value)	Logistic Regression Random Over-Sampling (Precision Value)	Logistic Regression Smote -Token (Precision Value)
Anchors	0.907	0.547	0.747
Lime	0.872	0.918	0.676
SHAP	0.891	0.924	0.752
BEEF	0.881	0.741	0.725
LORE	0.913	0.878	0.781

Conclusion

- One of the main relevant financial services is the credit risk assessment, whose aim is to support financial institutes in defining their policies and strategies.

- Predicting credit risk is a relevant challenge in the finance industry, particularly in Social Lending Platforms where high dimensionality and imbalanced data present unique challenges.

- Nevertheless, different challenges are faced by Social Lending Platform with respect to the traditional ones due to the high dimension and imbalanced data.

- This study proposes a benchmark for evaluating the effectiveness of machine learning techniques for credit risk prediction in real-world scenario, with a focus on managing imbalanced data sets and ensuring explainability.

Thank you for
your attention!