



College of Humanities
Master of Financial
Technology-
Fin Tech

Artificial Intelligence for Financial Technology

FNT 626

Lecture 05

Prepared By:

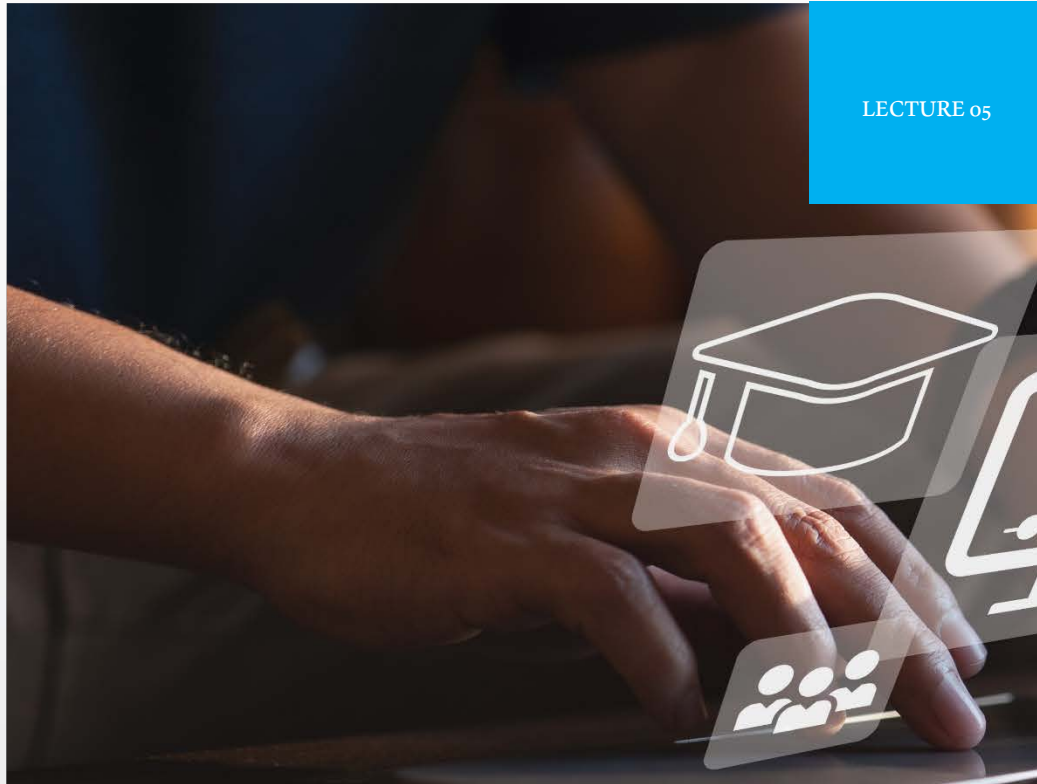
Dr. Abeir Moussa



Artificial Intelligence for Financial Technology

FNT 626- AI for Fin Tech

Lecture 05



Ch6. Machine learning, alternative data, and the regulation of consumer credit markets: The case of algorithmic credit scoring

Ch7. AI-based consumer credit underwriting-The role of a national credit database

LECTURE 05

Ch.6: Machine learning, alternative data, and the regulation of consumer credit markets

Chapter outline

1. INTRODUCTION
2. THE RISE OF ALGORITHMIC CREDIT SCORING
3. IMPACT OF ALGORITHMIC CREDIT SCORING ON CONSUMER CREDIT MARKETS
4. REGULATORY CHALLENGES AND OPPORTUNITIES
5. CONCLUSION

1. INTRODUCTION

In many countries, credit providers are required by law to assess the ‘creditworthiness’ of borrowers prior to extending credit. In the United Kingdom (UK), this requires an assessment of both the credit risk to the lender, as well as the affordability of the credit for the borrower. From a public regulatory perspective, the credit risk requirement seeks to support allocative efficiency in consumer credit markets – by ensuring that capital is allocated to the most valuable projects, and by minimizing adverse spillover effects (financial and real) due to systemic non-performing loans and to maximize profitability.

Ch.6: Machine learning, alternative data, and the regulation of consumer credit markets

2. THE RISE OF ALGORITHMIC CREDIT SCORING

Credit scoring entails the use of statistical techniques to model and predict a borrower's creditworthiness. Conventionally, credit scores have been calculated primarily using past credit performance data and, to a lesser extent, financial account transaction data and selected public data about borrowers, which are shared between credit providers through **credit reference agencies (CRAs)**. The emphasis on past credit performance reflects both the demonstrated statistical correlation between credit history and the credit risk of a borrower, as well as traditional limits on lenders' access to non-financial, and/or non-credit data about borrowers. As a result, these borrowers (often referred to as 'credit invisibles') are typically charged more to borrow, or denied credit by mainstream lenders, and forced to rely on 'payday lenders', or informal, unregulated sources of credit – often at punitive interest rates.

Ch.6: Machine learning, alternative data, and the regulation of consumer credit markets

2. THE RISE OF ALGORITHMIC CREDIT SCORING

FinTech lenders, such as Wonga and Zopa in the UK, and Lending Club, Prosper and Zest AI (formerly Zest Finance) in the United States (US), were pioneers of algorithmic credit scoring, partly due to lighter regulatory risk capital and modelling requirements compared with traditional lenders, such as banks. The first dimension relates to the use of **a larger volume and variety of data ('alternative' data)**. Regarding the first dimension, **alternative data** includes **both non-credit, financial data** (for example, rental, online subscription and mobile phone bill payments data), and **non-credit, non-financial data** (for example, 'social' data captured from consumers' social networks, and 'behavioral' data obtained by monitoring consumers' interaction with the lender's website or app). The second and more recent development embodied by algorithmic credit scoring is the use of ML techniques to analyze the data. ML is a sub-field of applied artificial intelligence (AI) research which involves training computer systems to learn patterns in data and apply their 'knowledge' to predict outcomes in previously unseen data. Notably, ML can more accurately capture non-linear relationships in the data, as well as reflect changes in the population and environment by 'learning' from new training data.

Ch.6: Machine learning, alternative data, and the regulation of consumer credit markets**3. IMPACT OF ALGORITHMIC CREDIT SCORING ON CONSUMER CREDIT MARKETS**

Expanding the number and types of measured variables, and employing ML techniques, allows for a more detailed, multi-dimensional observation of a borrower's characteristics that can be used to estimate their creditworthiness. This is especially useful for credit invisibles, who may present an acceptable credit risk despite not having any conventional, financial credit data to support this assessment. Many of these variables proxy for traditional credit lifecycle variables, such as income, whilst others proxy for information that is harder to observe or measure, such as the borrower's level of financial sophistication. By enabling a more accurate creditworthiness assessment in this way, algorithmic credit scoring could enhance allocative efficiency in consumer credit markets.

Furthermore, by facilitating better access to credit for credit invisibles, algorithmic credit scoring could help to reduce extant distributional unfairness, given that many of these borrowers come from low-income backgrounds. The automation of credit decisions using algorithmic credit scoring also offers to reduce inefficiency and unfairness due to 'taste-based' discrimination in lending, by substituting prejudiced (sexist or racist) loan officers with more objective, less prejudiced algorithms.

Ch.6: Machine learning, alternative data, and the regulation of consumer credit markets**3. IMPACT OF ALGORITHMIC CREDIT SCORING ON CONSUMER CREDIT MARKETS**

Conversely, the use of ML increases the scope for discrimination (and unfairness) due to reliance on proxies for protected characteristics. Discrimination and unfairness could also result from biases in the training, validation and/ or test datasets used to build ML scoring models. Additionally, the opacity and complexity of certain ML techniques make it more difficult to preempt ex ante or verify ex post the exploitation of, or unlawful discrimination against, borrowers. Relatedly, algorithmic credit scoring could become a source of both inefficiency and distributional unfairness in consumer credit markets by enabling the more effective exploitation of borrowers' cognitive and behavioral weaknesses using data-driven insights. The question arises whether these same technologies in the hands of borrowers could attenuate the risks due to their use by lenders. For example, ML and alternative data are already being used to build highly personalized web plugins and mobile apps that seek to counteract behavioral biases, and 'nudge' consumers into making better financial decisions.

Ch.6: Machine learning, alternative data, and the regulation of consumer credit markets**3. IMPACT OF ALGORITHMIC CREDIT SCORING ON CONSUMER CREDIT MARKETS**

However, the efficacy of these applications depends on the extent to which they are adopted by consumers. Moreover, to the extent that these solutions do not fully replace financial decision-making by consumers, consumers who are compromised by behavioral and cognitive weaknesses could simply ignore the advice offered by the relevant app. In particular, incumbent credit providers enjoy privileged access to aggregate financial transaction data and product use patterns gleaned from multiple transactions with borrowers over time, which could be difficult for third party consumer-helping platforms to substitute. The Tech-Fins – big tech companies such as Amazon and Apple that are rapidly expanding their consumer financial services offerings – have a considerable information advantage due to the vast volumes of proprietary data collected from their customers. However, the Open Banking platform has helped to improve access to consumer financial data for third party account information service providers.

Ch.6: Machine learning, alternative data, and the regulation of consumer credit markets**4. REGULATORY CHALLENGES AND OPPORTUNITIES**

Algorithmic credit scoring thus presents itself as a double-edged sword. On the one hand, it offers to improve allocative efficiency and distributional fairness in consumer credit markets, inter alia, by improving the accuracy of creditworthiness assessment and widening access to credit for credit invisible consumers. As a general matter, the existing principles and conduct-based approach of the UK consumer credit regulatory regime gives regulators the flexibility to respond dynamically to the use of new and fast-evolving technologies by market participants. In particular, the principles that firms must ‘treat customers fairly’ and act with ‘due care, skill and diligence’ provide a broad legal basis for regulators to respond to potential exploitation of consumers through the use of algorithmic credit scoring. Moreover, the existing affordability rules, as well as interest rate restrictions on ‘high-cost, short-term credit’, limit the scope for data-driven price discrimination and exploitation of vulnerable borrowers, at least in regulated consumer credit markets. To complement the exchange of information with firms under the principles-based approach, regulators could make greater use of ML techniques to detect, understand and remedy undesirable behavior by market participants more directly.

Ch.6: Machine learning, alternative data, and the regulation of consumer credit markets**4. REGULATORY CHALLENGES AND OPPORTUNITIES**

These findings can be used to inform regulatory changes, for example, mandating greater personalization of information disclosure by firms, or requiring firms to adjust the consumer choice architecture in a more targeted way (for example, changing the default settings on their website or app to mitigate common consumer mistakes). Relatedly, firms (banks and non-banks) could be required to put in place more robust governance and oversight arrangements specifically relating to their ML systems, including algorithmic credit scoring systems. Inter alia, this could encompass more specific procedures for data quality verification, as well as continuous model feedback testing, cross-validation, and auditing. These procedures should build on the auditing, certification, impact assessment and 'data protection by design and default' provisions under the GDPR and UK Data Protection Act 2018, as well as existing model risk management rules applicable to banks under the Basel Framework.

Ch.6: Machine learning, alternative data, and the regulation of consumer credit markets

4. REGULATORY CHALLENGES AND OPPORTUNITIES

Indeed, cross-sectoral data protection regulation provides an important additional mechanism for mitigating discriminatory and unfair treatment of credit consumers due to the processing of their personal data in algorithmic credit scoring. This includes the principles of ‘purpose limitation’ – requiring personal data to be collected only for ‘specified, explicit, and legitimate purposes and not further processed in a manner that is incompatible with those purposes’ – and ‘data minimization’ – requiring personal data to be ‘adequate, relevant and limited to what is necessary in relation to the purpose for which they are processed’. Furthermore, the GDPR expands the rights of data subjects to control the use of their data, including a potentially broader right to receive ‘meaningful information about the logic involved’ in automated decision-making (the so-called ‘right to explanation’). An expansive interpretation of this right (and corresponding duty) by credit providers and/or regulators could better support borrowers in challenging inaccurate and/or unfair credit decisions. Of course, there is a delicate balance between supporting and harming consumers through stricter enforcement of data protection laws. For example, a stricter explanation duty could also reduce the potential benefits of efficiency and fairness from algorithmic credit scoring if firms are restricted to using statistical techniques that are simpler and more ‘explainable’, yet less effective for assessing creditworthiness.

LECTURE 05

Ch.7: AI-based consumer credit underwriting, The role of a national credit database

Chapter outline

1. INTRODUCTION
2. AI-BASED CONSUMER CREDIT UNDERWRITING: OPPORTUNITIES, RISKS AND CHALLENGES
3. DATA ACCESS IN AI-BASED UNDERWRITING—CREDIT DATA-SHARING REGIMES
4. A HYBRID CREDIT DATABASE MODEL—POTENTIAL IMPLICATIONS FOR AI-BASED CREDIT UNDERWRITING
5. CONCLUSION

Ch.7: AI-based consumer credit underwriting- The role of a national credit database

1. INTRODUCTION

Credit underwriting—namely, the decision whether to provide credit to a consumer and under which conditions—depends on evaluating the default risk entailed in the provision of such credit. However, following the information revolution and the development of advanced automated technologies for transmitting and analyzing data, more and more credit providers have turned to rely on algorithmic credit underwriting, sometimes even without any human involvement. Today, lenders predict consumers' credit risk using algorithmic tools, such as machine learning (ML), that process vast amounts of diverse information quickly and efficiently. Lenders' turn to algorithmic underwriting was initially facilitated by credit data-sharing regimes, originally established in the United States and which have now become prevalent in many other countries as well. Over the years, these regimes have developed into an established industry, leading to the creation of a financial information market and an institutionalized framework for trading in financial information about individuals. A key player in credit data-sharing regimes is **the credit bureau**—a commercial entity that collects, processes and sells credit data and various services based on this data.

Ch.7: AI-based consumer credit underwriting, The role of a national credit database**1. INTRODUCTION**

‘Credit scoring’ stands at the core of these services. Many lenders purchase credit scores from credit bureaus as a key component in their underwriting process. These companies analyze vast amounts of data, including alternative information from various sources, such as videos, photos and audio, to find reliable indications of credit risk. Some even use information stored on smart devices that operate ‘Internet of Things’ technology to generate data that may help with assessing credit risk. Today, algorithmic credit underwriting goes far beyond traditional credit scoring by credit bureaus. In addition to purchasing external credit scoring services, many lenders apply internal AI-based models to predict credit risk. To build these models, they independently collect and analyze a wide range of data that may predict the creditworthiness or repayment ability of the customer. This data may rely on financial data records about their customers, possibly obtained through open banking systems, as well as non-financial data collected and analyzed by AI technologies, such as data about activity on social networks, frequency of use of smartphones, fitness, surfing activity on websites, and more. The expansion of the scope and type of data used to produce AI-based credit underwriting models has enabled drawing a considerably more accurate picture regarding the customer’s ability to repay their debts. In light of the growing potential of relying on non-financial data in consumer credit underwriting models, one may question the necessity of credit data-sharing mechanisms of the kind that exist in different countries.

2. AI-BASED CONSUMER CREDIT UNDERWRITING: OPPORTUNITIES, RISKS AND CHALLENGES

**2.1 From Manual to AI-based Consumer Credit
Underwriting**

2.2 The Opportunities of AI-based Underwriting

2.3 The Risks of AI-based Consumer Underwriting

LECTURE 05



**Ch.7: AI-based consumer credit underwriting,
The role of a national credit database**

Ch.7: AI-based consumer credit underwriting, The role of a national credit database

2. AI-BASED CONSUMER CREDIT UNDERWRITING: OPPORTUNITIES, RISKS AND CHALLENGES

2.1 From Manual to AI-based Consumer Credit Underwriting

Indeed, the accuracy and efficiency of predictive models depends on the four Vs of the input data: **volume, variety, velocity and veracity**. Lenders, hence, leverage diverse sources of data and sophisticated ML techniques to develop algorithmic models of credit underwriting. Unlike earlier algorithmic models for assessing creditworthiness, today's models use both financial data as well as non-financial (or 'alternative') data as their input. Credit scoring refers to 'the use of statistical techniques to classify and rank consumers according to their credit risk, as estimated from patterns in the past credit performance and financial account transaction data of consumers with similar attributes.' credit scores could be purchased from authorized credit bureaus or from private FinTech companies who specialize in providing credit risk assessments, depending on local financial regulations. Nevertheless, consumers' credit scores are not necessarily the definitive parameter in lenders' credit underwriting models. Rather, lenders usually combine this numerical indicator with additional financial data extracted from the consumer's credit history as documented by the lender, financial data available through systems of open banking, and alternative data available on public databases. Figure 7.1 demonstrates the different input data weighted by credit underwriting models.

Ch.7: AI-based consumer credit underwriting, The role of a national credit database

2.1 From Manual to AI-based Consumer Credit Underwriting

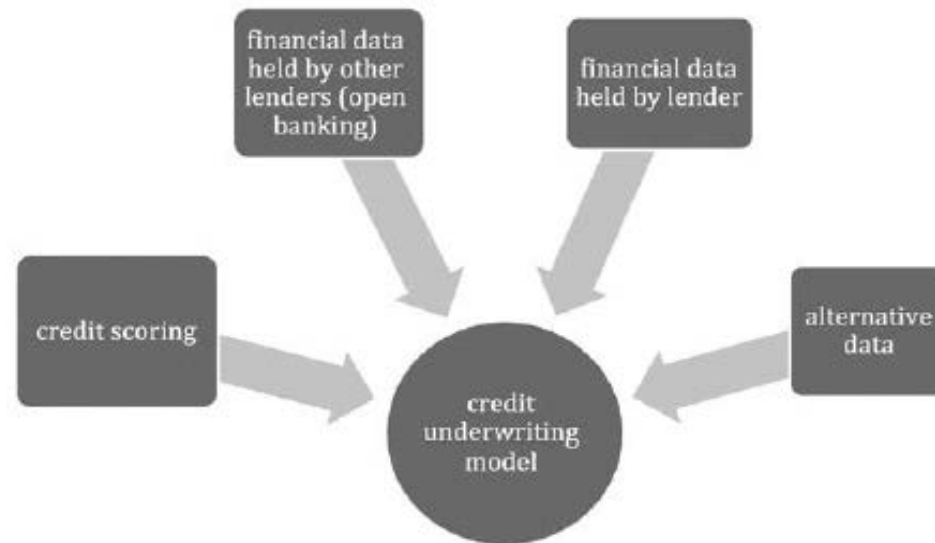


Figure 7.1 Input data weighted by credit underwriting models

Ch.7: AI-based consumer credit underwriting, The role of a national credit database**2. AI-BASED CONSUMER CREDIT UNDERWRITING: OPPORTUNITIES, RISKS AND CHALLENGES****2.1 From Manual to AI-based Consumer Credit Underwriting**

The shift to AI-based consumer credit underwriting offers an unprecedented opportunity to make prompt and accurate decisions about credit eligibility and pricing, and may open up the lending market to those who have once been credit invisible. At the same time, however, like other AI-based decision-making mechanisms, it may risk fundamental rights. These opportunities and risks are summarized below.

Ch.7: AI-based consumer credit underwriting, The role of a national credit database

2.2 The Opportunities of AI-based Underwriting

To begin with, the capabilities of AI make the process of consumer credit underwriting more efficient. Empirical evidence provides ample support to this proposition. The application of algorithmic decision-making in credit scoring reduces both the cost and the time of acquiring the relevant data capable of assessing the consumer's creditworthiness. Big data techniques extract, sort and analyze the relevant data in a speedy manner, replacing the slow and costly process of meeting consumers in person and documenting their financial information manually. Moreover, the use of predictive models that are based on ML, such as random forest, generally improve the accuracy of predicting credit default using simple logistic models. ML can capture unexpected correlations in data, such as the relation of mobile phone usage, health fitness activity or retail and online borrowing data to credit risk.

Additionally, since ML models process a much larger number of data points regarding the consumer, the relational weight given to each factor reduces, and consequently also the risk of error. Further and related, the use of AI for credit underwriting may promote competition in the lending industry, consequently expanding credit eligibility and increasing financial inclusion. Some groups in the population coming from low-income, less educated and ethnic minority backgrounds have systematically suffered from very low credit scores, not because of their payback potential, but because of their lacking financial history ('credit invisibles'). Indeed, in 2019, the American Consumer Financial Protection Bureau (CFPB) found that the application of ML to alternative data resulted in a rise of 27% in credit approval rates.

Ch.7: AI-based consumer credit underwriting, The role of a national credit database**2.3 The Risks of AI-based Consumer Underwriting**

Notwithstanding the important opportunities of using AI and ML in credit underwriting, it raises serious risks to fundamental rights that cannot be ignored. These sophisticated systems are all vulnerable to errors and bias stemming from the quality of either the input data which feeds the ML model, the model applied to analyze the data or both. The discriminatory impact of algorithmic credit scoring was revealed in relation to minorities, women and the LGBTQ community. Additionally, the exploitation of personal data to develop systems of algorithmic decision-making may risk consumers' privacy and autonomy. This is especially true given the growing use of alternative data in credit underwriting models. Of course, to address this challenge, regulators must take into account different factors, including the regional characteristics of the lending market, the level of AI innovation in the jurisdiction, the legal framework applicable to AI-based decision-making, as well as regulatory restrictions on data sharing and data access.

3. DATA ACCESS IN AI-BASED UNDERWRITING—CREDIT DATA SHARING REGIMES

3.1 The US Regime

3.2 The Israeli Model



LECTURE 05

**Ch.7: AI-based consumer credit underwriting,
The role of a national credit database**

Ch.7: AI-based consumer credit underwriting, The role of a national credit database

3. DATA ACCESS IN AI-BASED UNDERWRITING—CREDIT DATA-SHARING REGIMES

3.1 The US Regime

The American regime is based on **credit reporting agencies (CRAs)** as the main player in the credit data-sharing system. The CRAs gather and catalog data regarding the payment behavior of consumers and issue credit reports that include the cataloged information. The data includes both negative information indicating non-compliance with financial obligations and positive information demonstrating compliance. The main reporting sources of credit data are banks, credit card companies and other lenders. The information provided includes data on the customer's past and present credit accounts (account type, credit amount, balance, length of time the account has been defined as active, payment history), credit cards held by the customer, information on mortgage loans and credit facilities, and more. Negative information includes data on non-repayment of loans, data on returned cheques and payment instructions that have not been honored, debt claims, etc. Other sources of data reported to the CRAs are utility companies (electricity, water, gas), telecommunications companies, insurance companies (including medical insurance), property landlords, and large retail chains, to expand the scope of information about the customer and obtain as complete a picture as possible about their financial situation and reliability.

Ch.7: AI-based consumer credit underwriting, The role of a national credit database**3.1 The US Regime**

In addition to all of these, personal data such as medical information, residential history, education, and employment history are also furnished to the CRAs. Importantly, commercial institutions submit the above information to the CRAs voluntarily, on their own initiative; there is no legal obligation compelling its provision. Despite the absence of a reporting duty, the information sources submit information on billions of consumer credit accounts to one or more nationwide CRAs. The CRAs aggregate the data collected, prepare a credit report, and sell it to authorized third parties. The report usually includes information collected over the previous seven years; however, certain data, such as on bankruptcy, may be retained for ten years. The law prohibits a reporting institution to provide information when the institution 'knows or has reasonable cause to believe that the information is inaccurate', and it requires institutions to participate in the process of correcting errors that consumers identify. A multi-stage process is used to identify data irregularities and logical inconsistencies, both at the level of the general data file and at the level of the individual consumer's information. While American law obliges CRAs, when preparing a consumer report, to 'follow reasonable procedures to assure maximum possible accuracy' of the information, various studies have documented errors, inconsistencies, and outdated information in the credit reports.

Ch.7: AI-based consumer credit underwriting, The role of a national credit database**3.1 The US Regime**

In order to protect consumers, the law vests them with the right to access their credit reports, dispute the completeness or accuracy of their reports, ask for corrections, and annotate their records when resolutions are not achieved. Other creditors do not always report information on customers who consistently make their payments as scheduled (positive information). In addition to credit reports, CRAs provide ancillary services and products based on credit data, including credit scores. Moreover, many lenders develop and use their own proprietary scoring models based on the credit reports they received on their customers, eliminating the need to purchase scores from credit bureaus. Recently, with the increased use of AI for credit risk assessment, more and more FinTech companies have emerged offering lenders advanced algorithmic credit scoring models. Similar credit data regimes also exist in most European countries and in other jurisdictions: Commercial credit bureaus collect credit data from various sources of information, and sell the information as well as additional ancillary services such as credit scores to the lenders.

LECTURE 05

Ch.7: AI-based consumer credit underwriting, The role of a national credit database

3.2 The Israeli Model (Page 160)

4. A HYBRID CREDIT DATABASE MODEL- POTENTIAL IMPLICATIONS FOR AI-BASED CREDIT

UNDERWRITING

- 4.1 Efficiency**
- 4.2 Competition**
- 4.3 Accuracy**
- 4.4 Innovation**
- 4.5 Financial Inclusion**
- 4.6 Fairness**
- 4.7 Privacy and Autonomy**



LECTURE 05

Ch.7: AI-based consumer credit underwriting, The role of a national credit database

Ch.7: AI-based consumer credit underwriting, The role of a national credit database**4. A HYBRID CREDIT DATABASE MODEL—POTENTIAL IMPLICATIONS FOR AI-BASED CREDIT UNDERWRITING**

A hybrid data-sharing model relies on a national credit database with restricted access. Credit bureaus play a major role in this model, as they are the only entities who have access to the database

for developing scoring models. The hybrid model may implicate AI-based credit underwriting in different important ways. Recall that credit underwriting models used by lenders usually rely on algorithmic credit scores provided by credit bureaus, sometimes even exclusively. Hence, the surrounding data-sharing regime which supports the production of these scores may have direct implications for consumer credit underwriting models. Below we summarize these different implications.

Ch.7: AI-based consumer credit underwriting, The role of a national credit database**4.1 Efficiency**

The core benefit of using AI in credit underwriting relates to its efficiency. Under the hybrid model, lenders can keep applying AI techniques to develop their proprietary consumer underwriting models. Additionally, when all credit bureaus develop their credit scoring models on the basis of the same credit data held in the national database, they generate very similar results. For lenders, the fact that all credit scores generated by authorized credit bureaus are similar eliminates the need to combine different scores together.

Ch.7: AI-based consumer credit underwriting, The role of a national credit database**4.2 Competition**

On the one hand, a hybrid credit data-sharing model could promote competition, by assisting smaller lenders who cannot afford developing their own AI-based models, either because they lack a rich credit history of their own customers, or due to the lack of financial or technical resources needed for such development. At the same time, while dominant lenders can overcome this access barrier by deploying the rich credit history of their customers and develop their own algorithmic models, new lenders who lack such a rich database cannot. The Committee was of the opinion that providing lenders with direct access to the database may undermine the business model of the credit bureaus since lenders will become less dependent on their services. Note that allowing lenders direct access to the database may also have implications for customers' privacy.

Ch.7: AI-based consumer credit underwriting, The role of a national credit database**4.3 Accuracy**

The implementation of a hybrid credit data-sharing model should also accelerate the accuracy of the credit scores issued by credit bureaus and this, in turn, could improve the accuracy of the AI-based credit underwriting models developed and applied by lenders. First, the process of collecting and storing the credit data in the national database is strictly executed by the central bank, further assuring the reliability and correctness of the raw data. Second and related, scoring models developed by the credit bureaus are also subject to oversight by the central bank. Third, the transfer of credit data from reporting bodies to the national database is done mandatorily, further reducing the risk of partial or incomplete reporting, which subsequently reduces the accuracy of the scoring model as explained earlier. Fourth, the implementation of a hybrid model improves consumers' ability to inspect credit data reports for errors. Since their credit scores are solely derived from the financial data stored on the national database, it could become easier to understand the rational logic which underlies their scores. Further and related, since the national database only contains credit data, credit reports provided by credit bureaus, which offer no insight into alternative data, could become more intelligible to consumers.

Ch.7: AI-based consumer credit underwriting, The role of a national credit database**4.4 Innovation**

The hybrid model of credit data sharing may have different implications for innovation. On the one hand, the existence of multiple credit bureaus as a key player in the data-sharing regime should create competition between them which, in turn, would encourage them to adopt new technologies and offer innovative services in order to gain an advantage over competitors. Specifically, since all credit bureaus use the same information, the only way to differentiate themselves from their competitors is through the nature of the services they provide to credit providers, hence their incentive to use advanced technologies. On the other hand, the existence of a central national database as a single body for the collection and concentration of credit data, with no competitors in respect of these specific functions, may lead to conservatism and stagnation. In addition, the existence of a closed database, which includes only credit data, as opposed to other financial data and even non-financial data, may reduce innovation from the part of credit bureaus, because it denies access to other sorts of information which are vital for model development by credit providers and other market participants.

Ch.7: AI-based consumer credit underwriting, The role of a national credit database**4.5 Financial Inclusion**

Restricting the types of data used to develop credit scoring models to credit data only may undermine the inclusive potential of AI-based credit underwriting. Indeed, the increasing reliance on alternative data may expand credit eligibility and provide lending services to two groups of populations: credit invisibles, who lack sufficient financial history, and credit underserved, who are refused credit due to their very low credit scores. However, in the hybrid model, non-financial data is not included in the national database, and this may hurt the credit invisibles and the credit underserved. Hence, while some lenders (probably the dominant ones) might find it more efficient to mainly rely on rich credit data in their underwriting models, other lenders (probably innovative FinTech companies) may elect to leverage nontraditional data to reach the underserved. As for policymakers, it is possible to adopt different mechanisms that may offer credit to individuals with poor credit scores. One such mechanism could be a governmental fund that guarantees credit provided by designated lenders to consumers with very low credit scores.

Ch.7: AI-based consumer credit underwriting, The role of a national credit database

4.6 Fairness

Moving to the **risks** involved in deploying AI in consumer credit underwriting, it appears that implementing a central credit database could assist in mitigating them. When credit bureaus solely rely on strictly supervised financial data in developing their models, the risk of having biased input data is reduced. As Gillis explains, ‘algorithmic pricing could exacerbate the “biased world” problem because it increases the variance in predictions and may expand the number of “biased world” inputs through its use of nontraditional data.’ While lenders may still invest resources in extracting and analyzing alternative data to improve their internal underwriting models, it is nonetheless possible that due to the increased accuracy of the credit scores provided by credit bureaus, the costs of deploying alternative data will outweigh their benefits. Furthermore, when the data used to feed the credit scoring models is held and supervised in a central national database, it becomes much easier to inspect it and ensure it is not biased in itself or otherwise functions as a proxy of protected categories. Additionally, the scoring models themselves could be subject to ongoing oversight by the regulator to ensure they are fair and credible. To the extent that discriminatory inputs or processes are revealed, the regulator could demand and assure they are addressed properly.

Ch.7: AI-based consumer credit underwriting, The role of a national credit database**4.7 Privacy and Autonomy**

The hybrid model, which relies on a central, nationally supervised credit database, could contribute greatly to consumers' privacy and autonomy. First, as mentioned earlier, the data that is used by the credit bureaus to generate credit scores is unidentifiable. Even though data anonymization cannot perfectly protect consumers' privacy, it is obviously better than using identifiable information. Second, holding consumers' credit data at a nationally secured database is likely to reduce cyber-security risks that may result in information leakage or otherwise expose personal information to hacking. Third, if credit scores issued by authorized credit bureaus are exclusively based on financial data (and not on alternative data), consumers' ability to understand and control how the data is used to predict their creditworthiness greatly improves.

Conclusion

Algorithmic credit scoring, and the technologies underlying it, present both benefits and risks for consumer credit markets. Chapter six has argued that the broad principles- and conduct-based approach of UK consumer credit regulation provides the flexibility necessary for regulators and market participants to respond dynamically to these new technological risks. Supervisory authorities could also make greater use of ML and alternative data to strengthen the supervision of consumer credit firms. However, the interpretation of this regime in the consumer credit context needs to be carefully calibrated so as not to inhibit the potential benefits of new, data-driven technological applications such as algorithmic credit scoring.

Moreover, chapter seven explored how the implementation of a hybrid data-sharing regime, in which authorized credit bureaus enjoy exclusive access to financial data gathered and stored on a nationally secured database, may impact the benefits and risks of AI-based consumer underwriting.

Artificial Intelligence for Financial Technology

References:

Required Reference:

Remolina, N. and Gurrea-Martinez, A. eds., 2023. Artificial intelligence in finance: Challenges, opportunities and regulatory developments. Edward Elgar.

Recommended Readings

Chishti, S., 2020. The AI book: the artificial intelligence handbook for investors, entrepreneurs and fintech visionaries. John Wiley & Sons.

The End of Lecture

Thank You