



www.aileaders-project.eu

SYMULACJA – Wyzwania etyczne w analizie finansowej



Co-funded by
the European Union

Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor the entity providing the grant can be held responsible for them.

SYMULACJA

- Wyzwania etyczne w analizie finansowej

- 01  Streszczenie 3
- 02  Wprowadzenie 4
- 03  Prezentacja narzędzi 5
- 04  Wykonanie symulacji 6
- 05  Wnioski 7
- 06  Bibliografia 8

• 01 Streszczenie



Rodzaj OER

Demonstracja/symulacja przy użyciu Google Colab (RandomForest vs FairGBM Classifier)

Cel/przeznaczenie

Zapewnienie studentom praktycznej i krytycznej analizy tego, w jaki sposób algorytmiczne podejmowanie decyzji w analizie finansowej i prognozowaniu może reprodukcować nierówności strukturalne poprzez porównanie wyników uzyskanych na podstawie wskaźników sprawiedliwości. Symulacja ta zachęca do refleksji nad etycznymi aspektami automatyzacji finansowej i promuje opracowywanie bardziej sprawiedliwych i integracyjnych modeli prognostycznych.

Oczekiwane efekty kształcenia

Pod koniec symulacji studenci będą potrafili:

- 01** Wykrywać i interpretować algorytmiczne tendencje w procesie zatwierdzania kredytów;
- 02** Poddać refleksję nad etycznymi implikacjami zautomatyzowanych decyzji finansowych.
- 03** Wykorzystywać wskaźniki sprawiedliwości do oceny wyników modelu;

Słowa kluczowe

- Uczenie maszynowe
- Klasyfikacja
- Stronniczość
- Sprawiedliwość

Sugerowane podejście metodologiczne

Nauka oparta na problemach

UWAGA

Aby zrozumieć i pracować z treścią tego OER, wymagana jest średniozaawansowana znajomość **programowania** w języku **Python**.

• 02 Wprowadzenie



Narzędzia do prognozowania finansowego oparte na sztucznej inteligencji są coraz częściej wykorzystywane do automatyzacji decyzji kredytowych, takich jak zatwierdzanie pożyczek i ocena ryzyka (patrz np. Chen, 2020; Dastile i in., 2020; Chen i in., 2024; Heß & Damásio, 2025). Technologie te oferują znaczne korzyści pod względem wydajności, spójności i skalowalności. Jednak w kontekście ukształtowanym historycznie przez nierówności strukturalne budzą one poważne obawy etyczne.



Główną kwestią jest potencjał tych systemów do powielania – a nawet wzmacniania – istniejących uprzedzeń zakorzenionych w danych finansowych. Algorytmy, które wydają się neutralne, mogą, gdy są szkolone na podstawie stronicznych lub niekompletnych zbiorów danych, generować niesprawiedliwe wyniki, które w nieproporcjonalny sposób wpływają na już zmarginalizowane lub wrażliwe grupy społeczne.

Szczególnie przekonującym przykładem są kobiety (Orser i in., 2006; Ongena i Popov, 2016; Moro i in., 2017; Beck i in., 2018; De Andrés i in., 2021, między innymi) oraz osoby czarnoskóre (Chatterji & Seamans, 2012), które nadal borykają się z wieloma barierami w dostępie do kredytów udzielanych przez tradycyjne instytucje finansowe z powodu krzyżujących się form dyskryminacji.

To ćwiczenie oparte na symulacji zachęca studentów do krytycznego zbadania powiązań między uczeniem maszynowym, prognozowaniem finansowym i sprawiedliwością. Korzystając z syntetycznego zbioru danych odzwierciedlającego realistyczne wnioski kredytowe, studenci zbadają, w jaki sposób cechy takie jak płeć, zawód, wykształcenie i stan cywilny wpływają na wyniki zatwierdzania kredytów. Zbiór danych został celowo skonstruowany tak, aby zakodować subtelne wzorce stroniczności, oferując praktyczną możliwość wykrywania, analizowania i łagodzenia dysproporcji poprzez stosowane eksperymenty.

Ćwiczenie to, oparte na szeroko pojętej literaturze dotyczącej integracji finansowej i luk w finansowaniu, stanowi dla studentów wyzwanie do refleksji nad tym, w jaki sposób systemy sztucznej inteligencji mogą wzmacniać lub przeciwdziałać systemowej niesprawiedliwości. Dzięki temu studenci nie tylko zdobędą kompetencje techniczne w zakresie audytu sprawiedliwości i ograniczania uprzedzeń, ale także rozwiną krytyczne zrozumienie odpowiedzialności etycznej związanej z projektowaniem i wdrażaniem sztucznej inteligencji w ekosystemach finansowych.



• 03 Prezentacja narzędzi



Podstawowy scenariusz przedstawiony w tej symulacji koncentruje się na zbiorze danych dotyczących syntetycznych wniosków kredytowych.

Dane obejmują cechy demograficzne i społeczno-ekonomiczne, takie jak

- Płeć
- Rasa
- Wiek
- Zawód
- Poziom wykształcenia
- Stan cywilny

*Zmienną docelową jest **zatwierdzenie kredytu**, modelowane jako problem klasyfikacji binarnej.*

Chociaż zbiór danych został skonstruowany tak, aby odzwierciedlał realistyczne warunki, celowo zawiera ukryte wzorce stronniczości. Na przykład kobiety-przedsiębiorcy z obszarów wiejskich mogą spotykać się z wyższym odsetkiem odmów niż ich koledzy lub kobiety z obszarów miejskich, nawet przy uwzględnieniu kwalifikacji i czynników ryzyka. Symuluje to, w jaki sposób historyczne i systemowe bariery mogą być powielane w systemach automatycznych.

Celem jest zbadanie dwóch podstawowych kwestii

- 01** Czy algorytmy uczenia maszynowego w nieproporcjonalny sposób „zalecają” odrzucenie wniosków kredytowych składanych przez osoby czarnoskóre i kobiety w binarnych modelach predykcyjnych?
- 02** Jakie czynniki strukturalne lub algorytmiczne mogą przyczyniać się do tej dysproporcji?

• 04 Wykonanie symulacji



01 Dostęp do zeszytu symulacji

Przejdź do strony <https://tinyurl.com/2txdk9s>

02 Uruchom kod

Uruchom kolejno wszystkie komórki w notatniku. Upewnij się, że nie występują żadne błędy i że wszystkie wyniki są wyświetlane poprawnie. W tym celu kliknij przycisk odtwarzania w lewym górnym rogu każdej komórki.

03 Zapoznaj się z zestawem danych¹

Przejrzyj strukturę i zawartość zbioru danych. Zwróć szczególną uwagę na kluczowe zmienne demograficzne, takie jak płeć, zawód, wykształcenie i wynik kredytu. Zwróć uwagę na wszelkie potencjalne nierówności lub wzorce, które mogą wskazywać na stronniczość.

04 Analiza sprawiedliwości

Przeanalizuj wyniki. Porównaj wskaźniki zatwierdzeń w różnych grupach demograficznych, zwracając szczególną uwagę na płeć i rasę. Wykorzystaj wskaźniki sprawiedliwości podane w notatniku (np. parytet demograficzny, nierówny wpływ), aby oszacować potencjalne stronniczości.

05 Porównaj wyniki

Dokładnie porównaj wyniki modelu przed i po zastosowaniu interwencji dotyczących sprawiedliwości, jeśli zostały one uwzględnione.

06 Zastanów się nad wynikami

Na podstawie zaobserwowanych wyników zastanów się nad etycznymi i praktycznymi implikacjami wykorzystania sztucznej inteligencji w podejmowaniu decyzji finansowych. Rozważ kompromisy między dokładnością, wydajnością i sprawiedliwością.

1. Samouczek dotyczący narzędzia What If — <https://www.youtube.com/watch?v=jHojeFCc5HE>



Dane, na których oparliśmy tę symulację, są szeroko wykorzystywane do badania sprawiedliwości i stronniczości w uczeniu maszynowym, ponieważ zawierają wrażliwe atrybuty, takie jak płeć i rasa. Są one powszechnie stosowane do przewidywania, czy dochód danej osoby przekracza 50 000 dolarów rocznie, na podstawie takich cech, jak wiek, wykształcenie, zawód i stan cywilny.

Nasza symulacja pokazuje, w jaki sposób systemy prognozowania finansowego oparte na sztucznej inteligencji, mogą utrzymywać strukturalne uprzedzenia zakorzenione w danych historycznych, jeśli nie zostaną poddane weryfikacji

Wynika z niej, że dokładność prognozowania odmowy udzielenia kredytu w modelach binarnych jest wyższa w przypadku osób czarnoskórych i kobiet niż w przypadku osób rasy białej i mężczyzn.

Podsumowanie wyników



Dokładność w przewidywaniu decyzji o odrzuceniu wniosku kredytowego — tj. prawdziwe „wyniki negatywne” (%)

Czarny

97
KOBIETA

91
MĘŻCZYŹNI

Biały

94
KOBIETY

86
MĘŻCZYŹNI

Nawet w uproszczonych modelach, takich jak ten, obecność stronniczych wyników — takich jak niższy odsetek wyników fałszywie ujemnych w przewidywanych wskaźnikach odrzucenia dla osób czarnoskórych i kobiet — podkreśla etyczną odpowiedzialność programistów i analityków za kontrolowanie i udoskonalanie systemów sztucznej inteligencji. Wyniki te ponownie rozpalają dyskusję na temat martwych punktów w modelach oceny zdolności kredytowej (patrz np. Robb & Robinson, 2018).

Angażując się bezpośrednio w pomiary sprawiedliwości i strategię łagodzenia skutków, studenci nie tylko zdobywają narzędzia techniczne do identyfikacji stronniczości, ale także rozwijają głębsze zrozumienie tego, jak takie stronniczości się przejawiają i jak można je eliminować. Ćwiczenie to wzmacnia zasadę, że sprawiedliwość w finansowej sztucznej inteligencji nie jest kwestią drugorzędną, ale podstawowym elementem odpowiedzialnych i zrównoważonych innowacji.

• 06 Bibliografia



- Beck, T., Behr, P., & Madestam, A. (2011). Sex and credit: Is there a gender bias in lending?. European Banking Center Discussion Paper, (2011-027).
- Chatterji, A. K., & Seamans, R. C. (2012). Entrepreneurial finance, credit cards, and race. *Journal of Financial Economics*, 106(1), 182-195.
- Chen, T. H. (2020). Do you know your customer? Bank risk assessment based on machine learning. *Applied Soft Computing*, 86, 105779.
- Chen, Y., Calabrese, R., & Martin-Barragan, B. (2024). Interpretable machine learning for imbalanced credit scoring datasets. *European Journal of Operational Research*, 312(1), 357-372.
- Dastile, X., Celik, T., & Potsane, M. (2020). Statistical and machine learning models in credit scoring: A systematic literature survey. *Applied Soft Computing*, 91, 106263.
- De Andrés, P., Gimeno, R., & de Cabo, R. M. (2021). The gender gap in bank credit access. *Journal of Corporate Finance*, 71, 101782.
- Heß, V. L., & Damásio, B. (2025). Machine learning in banking risk management: Mapping a decade of evolution. *International Journal of Information Management Data Insights*, 5(1), 100324.
- Moro, A., Wisniewski, T. P., & Mantovani, G. M. (2017). Does a manager's gender matter when accessing credit? Evidence from European data. *Journal of banking & finance*, 80, 119-134.
- Ongena, S., & Popov, A. (2016). Gender bias and credit access. *Journal of Money, Credit and Banking*, 48(8), 1691-1724.
- Orser, B. J., Riding, A. L., & Manley, K. (2006). Women entrepreneurs and financial capital. *Entrepreneurship Theory and practice*, 30(5), 643-665.
- Robb, A., & Robinson, D. T. (2018). Testing for racial bias in business credit scores. *Small Business Economics*, 50, 429-443.





leaders

Śledź naszą podróż



www.aileaders-project.eu



Co-funded by
the European Union

Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor the entity providing the grant can be held responsible for them.