



# Zastosowanie metod uczenia maszynowego w procesie oceny spółek kapitałowych

---

MGR ALEKSANDRA SZYMURA  
ALEKSANDRA.SZYMURA@UE.WROC.PL



Uniwersytet Ekonomiczny  
we Wrocławiu



# Agenda

---

1. Metody oceny podmiotów gospodarczych
2. Uczenie maszynowe w ocenie przedsiębiorstw – przegląd literatury
3. Koncepcja badań
4. Ocena zasadności przeprowadzanych badań
5. „Proof of concept”

# Istota oceny przedsiębiorstw

---

1. Dostarczanie kluczowych informacji o funkcjonowaniu przedsiębiorstwa
2. Weryfikacja potencjalnych klientów lub kontrahentów
3. Minimalizacja ryzyka przy podejmowaniu decyzji
4. Maksymalizacja zysków z działalności
5. Zapobieganie wystąpienia negatywnych zdarzeń wpływających na przedsiębiorstwo
6. Szybsza adaptacja do zmieniających się warunków w otoczeniu

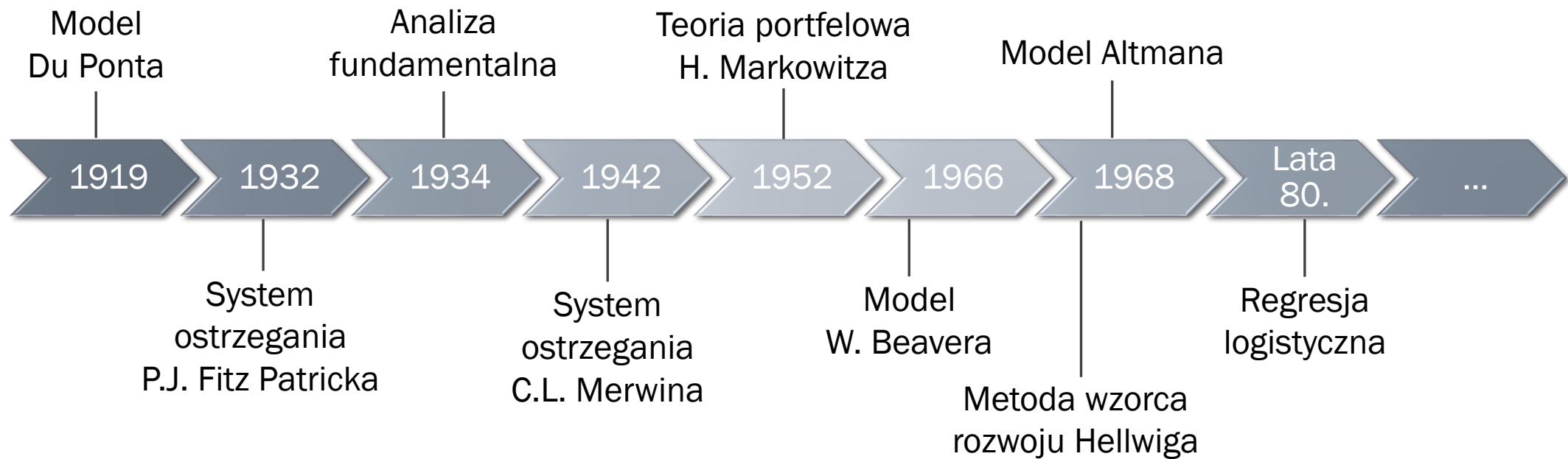
# Metody oceny przedsiębiorstw

---



# Zarys historyczny

---



# Uczenie maszynowe – definicja

---

*Machine Learning to zdolność komputerów do uczenia się bez programowania nowych umiejętności wprost.*

Arthur Samuel, 1959 rok

# Uczenie maszynowe

---



**Uczenie nadzorowane**



**Uczenie nienadzorowane**



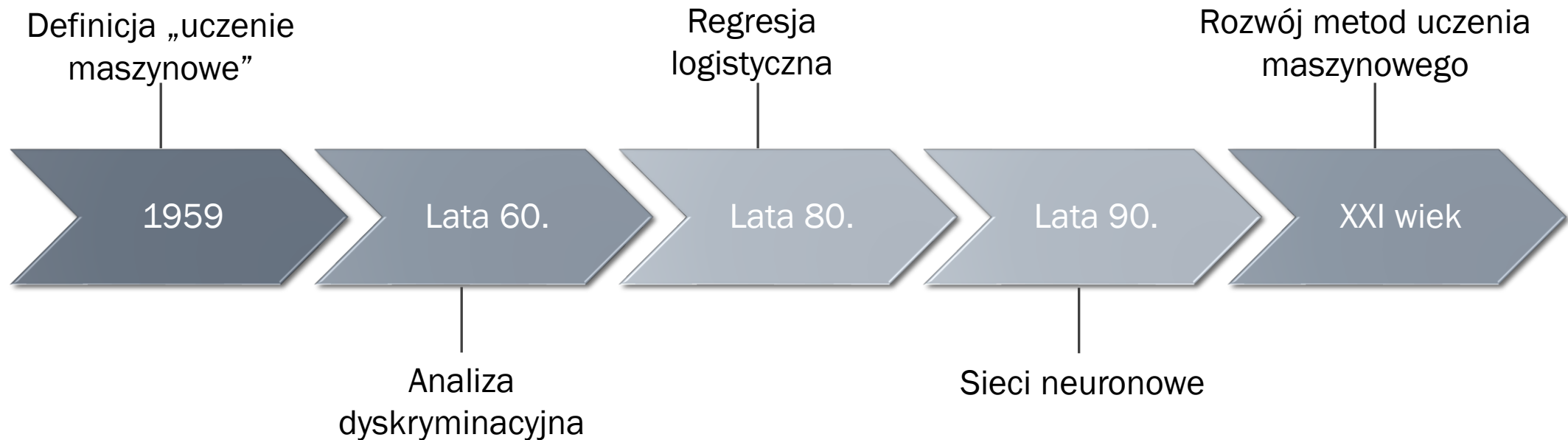
**Uczenie częściowo nadzorowane**



**Uczenie przez wzmacnianie**

# Zarys historyczny

---





# Przegląd literatury

---

1. Tsai, Wu (2008) – predykcja bankructwa oraz credit scoring w oparciu o sieci neuronowe. Wykorzystanie danych kredytowych pochodzących z Niemiec, Japonii i Australii.
2. Jabeur, Gharib i in. (2021) – zastosowanie m. in. metod opartych na wzmocnieniu gradientowym (Gradient Boosting, XGBoost, CatBoost) w predykcji niepowodzeń przedsiębiorstw w horyzoncie trzyletnim. Za każdym razem CatBoost okazał się najskuteczniejszy.
3. Pham, Ho (2021) – wykorzystanie algorytmów typu boosting (AdaBoost, XGBoost, Gradient Boosting) w predykcji upadłości banków. Uwzględniono wskaźniki finansowe.
4. Zhang, Wu i in. (2022) – predykcja złej sytuacji finansowej między innymi przy użyciu metod: XGBoost oraz LightGBM z uwzględnieniem wskaźników finansowych.

# Przegląd literatury

---

1. W badaniach przeważa podejście uwzględnienia jako zmiennej celu, czy dany podmiot ogłosił upadłość, czy nie lub, czy nadal istnieje, czy już jest nieaktywny
2. Z reguły w badaniach dokonywana jest predykcja sytuacji krańcowej, a często są pomijane aspekty prognozowania sytuacji, które do nich prowadzą np. utrata płynności, czy nadmierne zadłużenie
3. Ze względu na łatwą interpretowalność, bardzo często stosuje się klasyczne metody jak regresja logistyczna, czy drzewo decyzyjne, niekoniecznie cechujące się wysoką trafnością typowań. Dopiero w ostatnich kilku latach można zauważyć wzrost zainteresowania tzw. metodami czarnoskrzynkowymi, m. in. opartymi na wzmocnieniu gradientowym

# Koncepcja badań

---

1. Cel główny: zaproponowanie algorytmu pozwalającego na skuteczną klasyfikację podmiotu gospodarczego jako „ryzykowny” pod kątem wystąpienia wybranych zdarzeń lub nie
2. Wielokryterialna ocena spółek kapitałowych (spółki akcyjne oraz spółki z o. o.) – analiza ryzyka wystąpienia negatywnych zjawisk: utrata płynności, strata z działalności oraz nadmierne zadłużenie
3. Połączenie wykorzystania metod uczenia nienadzorowanego (eksploracja danych) wraz z technikami uczenia nadzorowanego (predykcja zdarzeń – klasyfikacja)
4. Zmienne oparte na wskaźnikach finansowych, ale również porównaniach r/r niektórych pozycji pochodzących ze sprawozdań (źródło: baza Orbis)
5. Wykluczenie z próby badawczej niektórych podmiotów jak np. banki – dla nich zasadne jest wykonanie analizy sektorowej

# Ocena zasadności badań

---

1. Szersze spojrzenie na sytuację finansową przedsiębiorstw pod kątem analizy wybranych rodzajów ryzyka działalności podmiotu
2. Spadek kosztów dla przedsiębiorstw w porównaniu z zastosowaniem innych metod np. z *due diligence*
3. Gromadzenie i przetwarzanie dużych ilości danych wymusza na przedsiębiorstwie dopasowanie się do otoczenia, szybsze reagowanie, automatyzację procesów
4. Analiza zjawisk wpływających na pogorszenie się kondycji finansowej przedsiębiorstwa – sygnał ostrzegawczy dla osób zarządzających, czy inwestorów

# Oryginalność pracy doktorskiej

---

1. Zapewnienie interpretowalności otrzymanych wyników końcowych
2. Podejście wieloetapowe w tworzeniu rozwiązania
3. Wielokryterialna ocena spółek kapitałowych – predykcja wystąpienia kilku zjawisk
4. Osobna ocena spółek akcyjnych oraz spółek z o. o. – możliwość porównania wyników dla obu typów spółek
5. Brak w polskich realiach badań uwzględniających jednocześnie podejście wieloetapowe oraz wielokryterialną ocenę podmiotów
6. Badanie zjawisk zwiększających ryzyko pogorszenia się sytuacji finansowej

# „Proof of concept”

---

1. Przygotowanie modeli dla próbki spółek akcyjnych bazując na wybranej puli zmiennych niezależnych opartych na wskaźnikach finansowych
2. Predykcja takich zjawisk jak: zapłata istotnej kary lub odszkodowania, utrata płynności, strata z działalności
3. Zmienne zależne ustalane na bazie obliczonych wartości progowych (zapłata istotnej kary), jak i wartości wskaźników uznanych za akceptowalne na bazie literatury (utrata płynności)

# „Proof of concept”

---

## 1. Metodyka badawcza:

1. Analiza zmienności oraz skośności cech
2. Analiza korelacji między zmiennymi
3. Podział zbioru na treningowy i testowy – 75/25
4. Równoważenie zbioru treningowego – technika SMOTE
5. Modelowanie: regresja logistyczna, drzewo decyzyjne, Gradient Boosting, XGBoost, LightGBM, CatBoost
6. SHapley Additive exPlanations

# „Proof of concept” – wyniki

---

Cecha	Predykcja kar i odszkodowań	Ryzyko osiągnięcia straty z działalności	Ryzyko utraty płynności
Skuteczność	CatBoost	Regresja logistyczna	LightGBM/CatBoost
Stabilność	Gradient Boosting/XGBoost	XGBoost	Drzewo decyzyjne

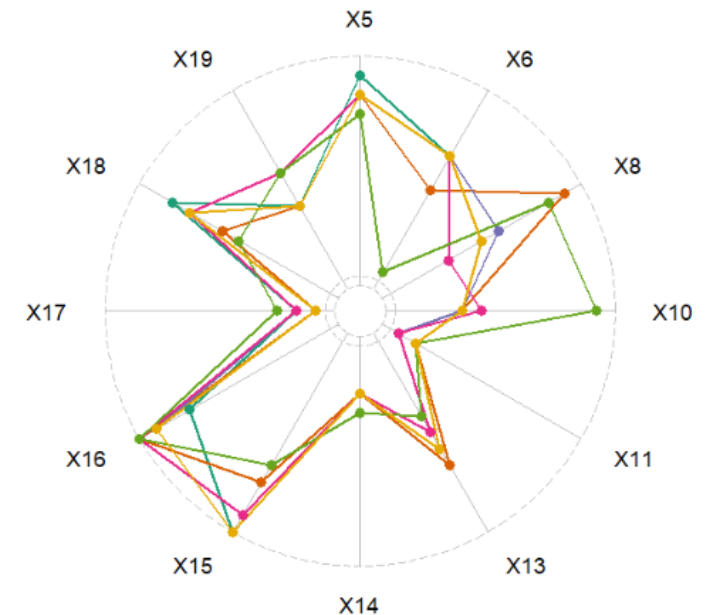


# „Proof of concept”

1. Zapewnienie interpretowalności globalnej otrzymanych wyników z wykorzystaniem techniki SHapley Additive exPlanations (SHAP)
2. Tworzenie rankingu ważności cech uwzględniając rezultaty wszystkich iteracji



Ranking cech według wartości Shapley'a



# Literatura

---

1. Jabeur, Sami Ben, Cheima Gharib, Salma Mefteh-Wali, and Wissal Ben Arfi. 2021. CatBoost model and artificial intelligence techniques for corporate failure prediction. *Technological Forecasting and Social Change* 166: 120658. <https://doi.org/10.1016/j.techfore.2021.120658>.
2. Pham, Xuan T.T., and Tin H. Ho. 2021. Using boosting algorithms to predict bank failure: An untold story. *International Review of Economics & Finance* 76: 40–54. <https://doi.org/10.1016/j.iref.2021.05.005>.
3. Szymura, Aleksandra. 2022. Risk Assessment of Polish Joint Stock Companies: Prediction of Penalties or Compensation Payments. *Risks*. <https://doi.org/10.3390/risks10050102>.
4. Tsai, C, and J Wu. 2008. Using neural network ensembles for bankruptcy prediction and credit scoring. *Expert Systems with Applications* 34: 2639–49. <https://doi.org/10.1016/j.eswa.2007.05.019>.
5. Zhang, Zijiao, Chong Wu, Shiyu Qu, and Xiaofang Chen. 2022. An explainable artificial intelligence approach for financial distress prediction. *Information Processing & Management* 59: 102988. <https://doi.org/10.1016/j.ipm.2022.102988>.

Dziękuję za uwagę! 😊

---