









Statistica piețelor financiare

Capitolul 12: Modele de scoring



Daniel Traian PELE

Academia de Studii Economice din București
IDA Institute Digital Assets
Blockchain Research Center
AI4EFin Artificial Intelligence for Energy Finance
Academia Română, Institutul de Prognoză Economică
MSCA Digital Finance

Acronime folosite în acest material (1/2)

- **AI**: Artificial Intelligence (inteligență artificială)
- **AR**: Accuracy Ratio (of the CAP curve) — raportul de acurateţe (al curbei CAP)
- **AUC**: Area Under the (ROC) Curve (aria de sub curba ROC)
- **BIS**: Bank for International Settlements (Banca Reglementelor Internaţionale)
- **BS**: Brier Score (scorul Brier)
- **CAP**: Cumulative Accuracy Profile (profilul cumulat al acurateţei)
- **CDS**: Credit Default Swap (swap pe riscul de nerambursare)
- **CI**: Confidence Interval (interval de încredere)
- **DD**: Distance to Default (distanţa pînă la nerambursare)
- **DM**: Deutsche Mark (marca germană)
- **EAD**: Exposure At Default (expunerea la momentul nerambursării)
- **EBIT**: Earnings Before Interest and Taxes (profitul înainte de dobînzi şi impozite)
- **FICO**: Fair Isaac Corporation — compania Fair Isaac (scorul FICO)
- **FN**: False Negative (fals negativ)
- **FPR**: False Positive Rate (rata fals pozitivelor)
- **GBM**: Geometric Brownian Motion (mişcarea browniană geometrică)
- **HL**: Hosmer–Lemeshow (test) — testul Hosmer–Lemeshow
- **IRB**: Internal Ratings-Based (approach) — abordarea pe baza ratingurilor interne
- **KS**: Kolmogorov–Smirnov (statistic) — statistica Kolmogorov–Smirnov

Acronime folosite în acest material (2/2)

- **LDA**: Linear Discriminant Analysis (analiza discriminantă liniară)
- **LGD**: Loss Given Default (pierderea în caz de nerambursare)
- **LR**: Likelihood Ratio (test) — testul raportului de verosimilitate
- **MVA**: Applied Multivariate Statistical Analysis (Härdle and Simar) — manualul de analiză statistică multivariată al lui Härdle și Simar
- **NYU**: New York University (Universitatea din New York)
- **PD**: Probability of Default (probabilitatea de nerambursare)
- **PDO**: Points to Double the Odds (punctele care dublează șansa bun:rău)
- **ROC**: Receiver Operating Characteristic (curve) — curba caracteristică de operare
- **RWA**: Risk-Weighted Assets (activele ponderate la risc)
- **SFE**: Statistics of Financial Markets, Exercises (Quantlet collection of the textbook) — colecția de Quantlet-uri a manualului Statistics of Financial Markets
- **SUA**: Statele Unite ale Americii
- **TA**: Total Assets (activele totale)
- **TN**: True Negative (adevărat negativ)
- **TP**: True Positive (adevărat pozitiv)
- **TPR**: True Positive Rate — rata adevărat pozitivelor (sensibilitatea)
- **UCI**: University of California, Irvine (Machine Learning Repository) — Universitatea din California, Irvine (arhiva de date pentru învățare automată)
- **UE**: Uniunea Europeană

Întrebarea de azi și traseul

- **Întrebarea**: cum transformă o bancă datele unui solicitant într-o probabilitate de nerambursare și de unde știm că cifra este corectă?
  - Capitolele 2–10 au măsurat riscul prețurilor; acest capitol măsoară riscul ca un debitor să nu-și ramburseze creditul
- **Traseul** capitolului
  - scorurile de credit, PD, LGD, EAD și capitalul Basel; datele de nerambursare și dezechilibrul claselor
  - Weight of Evidence și Information Value; regresia logistică și analiza discriminantă a lui Fisher
  - scorecard-ul și validarea lui: matricea de confuzie, ROC, AUC, Gini, CAP, KS, scorul Brier, calibrarea
  - reject inference și echitatea (fairness); scorul Z al lui Altman și distanța pînă la nerambursare a lui Merton

Rezultatele învățării

- Definiți PD, LGD, EAD și pierderea așteptată și explicați ce adaugă capitalul IRB din Basel peste pierderea așteptată
- Calculați Weight of Evidence și Information Value ale unei variabile dintr-un tabel de frecvențe
- Estimați o regresie logistică și interpretați coeficienții ca modificări ale logaritmului șansei și ca rapoarte ale șanselor
- Deduceți discriminantul liniar al lui Fisher și comparați-l cu modelul logit
- Scalați un scorecard și validați-l în afara eșantionului cu matricea de confuzie, AUC, Gini, KS, scorul Brier și graficul de calibrare

Bibliografie și instrumente

- Manual: Franke, Härdle & Hafner (2019), cap. 21 (probabilitatea de nerambursare) și cap. 22 (managementul riscului de credit)
  - analiza discriminantă: Härdle & Simar (2019), cap. 14
  - scoring de credit: Thomas, Crook & Edelman (2017); scorecard-uri: Siddiqi (2006)
- Quantlet-urile Python ale capitolului: Quantlets/Ch_12
  - PD condiționată din modelul cu un factor este portată din Quantlet-ul SFE SFEdefaproba
- Notebook-ul cursului: deschideți în Google Colab
- Curs video: MVA Multivariate Statistical Analysis

Decizia de creditare

- Un solicitant cere un credit; banca trebuie să decidă **acceptarea** sau **respingerea** și dobînda
  - **nerambursarea** (default): debitorul nu plătește conform contractului; Basel: banca apreciază că debitorul nu va plăti sau debitorul întîrzie peste 90 de zile o obligație semnificativă Comitetul de la Basel (2006), ¶452
- **Scorul de credit**: un număr calculat din datele solicitantului, care ordonează solicitanții după riscul lor de nerambursare
  - **PD** (probability of default): probabilitatea de nerambursare pe un orizont, de obicei un an
  - un scor bun **ordonează** corect (discriminare) și dă **nivelul** corect al PD (calibrare)
- Tipuri de scoruri
  - scoruri de aplicare (clienți noi) și scoruri comportamentale (clienți existenți, din istoricul plăților)
  - scoruri pentru persoane fizice (modele statistice pe multe credite mici) și pentru companii (indicatori contabili, prețuri de piață)

Începuturile scoringului de credit: Durand și FICO

- 1941: Durand (1941) separă creditele bune de cele rele ale creditorilor americani cu analiza discriminantă a lui Fisher
  - primul studiu statistic de scoring de credit
- 1956: Bill Fair și Earl Isaac înființează Fair, Isaac and Company (azi FICO) myFICO, istoria scorului FICO
  - 1989: scorul FICO, calculat din rapoartele birourilor de credit, între 300 și 850 de puncte
- Scoringul statistic a înlocuit judecata ofițerului de credit cu o regulă identică pentru toți Hand & Henley (1997)
  - mai ieftin, mai rapid și verificabil; rămîne însă întrebarea dacă regula este echitabilă



Sediul FICO, San Jose, California

Foto: Coolcaesar (2020); CC BY-SA 4.0; Wikimedia Commons

PD, LGD, EAD și pierderea așteptată

- Trei parametri de risc ai unui credit (Basel II, Comitetul de la Basel (2006))
  - **PD**: probabilitatea de nerambursare într-un an
  - **LGD** (loss given default): partea din expunere pierdută în caz de nerambursare, după recuperări
  - **EAD** (exposure at default): suma datorată în momentul nerambursării
- **Pierderea așteptată**: $\mathrm{EL} = \mathrm{PD} \times \mathrm{LGD} \times \mathrm{EAD}$ (cu PD și LGD independente)
  - **Exemplu**: un credit de 100 000 de lei cu PD $= 2\%$, LGD $= 45\%$: EL $= 0{,}02 \times 0{,}45 \times 100\,000 = 900$ de lei
  - EL este un cost al activității: intră în dobîndă și este acoperită prin provizioane
- Capitalul acoperă **pierderea neașteptată**: pierderile de peste EL într-un an prost

Formula de capital IRB din Basel

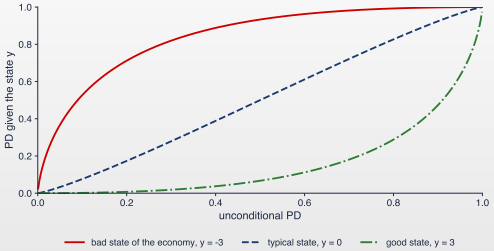
- Abordarea **IRB** (internal ratings-based): banca estimează PD (și eventual LGD, EAD); autoritatea de reglementare dă formula Comitetul de la Basel (2006), ¶328–330
  - $K = \mathrm{LGD}\left[\Phi\left(\frac{\Phi^{-1}(\mathrm{PD}) + \sqrt{R}\,\Phi^{-1}(0{,}999)}{\sqrt{1-R}}\right) - \mathrm{PD}\right]$; RWA $= 12{,}5\,K\,\mathrm{EAD}$
  - K: capitalul pe unitatea de EAD; Φ, Φ^{-1}: funcția de repartiție $N(0,1)$ și inversa ei; 0,999: un an prost apare o dată la 1000 de ani
  - R: corelația debitorilor cu un factor comun; RWA (risk-weighted assets): activele ponderate la risc, din care banca deține cel puțin 8%
- Logica Gordy (2003): termenul din $\Phi(\cdot)$ este PD într-un an al cărui factor comun este depășit doar cu probabilitatea 0,1%
  - deci K este VaR 0,1% al pierderii din credit minus EL: pierderea neașteptată (Capitolul 10)



Turnul BIS din Basel

Foto: Taxiarchos228 (2014); CC BY-SA 4.0; Wikimedia Commons

PD în anii favorabili și nefavorabili: modelul cu un factor



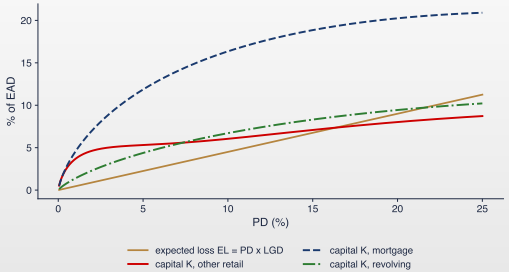
- Debitorul i intră în nerambursare dacă $\sqrt{\rho}\,Y + \sqrt{1-\rho}\,\varepsilon_i < \Phi^{-1}(\mathrm{PD})$; Y: starea economiei; dat fiind $Y = y$: $\mathrm{PD}(y) = \Phi\big((\Phi^{-1}(\mathrm{PD}) - \sqrt{\rho}\,y)/\sqrt{1-\rho}\big)$, $\rho = 0{,}2$
- ε_i: șocul propriu al debitorului; Y și ε_i independente, $N(0,1)$; ρ: ponderea factorului comun
- Interpretare: un credit cu PD $= 2\%$ intră în nerambursare cu probabilitatea 21,3% într-un an prost ($y = -3$), 1,1% într-un an obișnuit și aproape niciodată într-un an bun; IRB folosește $y = \Phi^{-1}(0{,}001)$

SFM_ch12_validation_irb

Exemplu rezolvat: capitalul IRB al unui credit de retail

- Alte expuneri de retail: $R = 0{,}03\,w + 0{,}16\,(1-w)$, $w = (1-e^{-35\,\mathrm{PD}})/(1-e^{-35})$ Comitetul de la Basel (2006), ¶330
  - $w \in [0,1]$ crește cu PD, deci R scade de la 0,16 (credite foarte sigure) la 0,03 (credite foarte riscante)
  - PD = 2%: $R = 0{,}095$; $\Phi^{-1}(0{,}02) = -2{,}054$, $\Phi^{-1}(0{,}999) = 3{,}090$
- **Pas cu pas**
  - argumentul: $(-2{,}054 + \sqrt{0{,}095} \times 3{,}090)/\sqrt{1-0{,}095} = -1{,}160$; $\Phi(-1{,}160) = 12{,}31\%$: PD într-un an prost
  - $K = 0{,}45 \times (12{,}31\% - 2\%) = 4{,}64\%$ din EAD: 4 639 de lei pentru un credit de 100 000 de lei
  - RWA $= 12{,}5 \times 4\,639 = 57\,986$ de lei: o pondere de risc de 58%
- Interpretare: EL (900 de lei) intră în preț; capitalul (4 639 de lei) absoarbe un an în care nerambursările ajung la 12,31% în loc de 2%

Pierderea așteptată și capitalul IRB în funcție de PD



- LGD $= 45\%$; capitalul K pentru alte expuneri de retail (R de la 0,16 la 0,03), credite ipotecare ($R = 0{,}15$) și expuneri revolving ($R = 0{,}04$)
- Interpretare: EL crește liniar cu PD; K crește repede la PD mici și apoi se aplatizează, deoarece un credit foarte riscant este în mare parte o pierdere așteptată (inclusă în preț)

SFM_ch12_validation_irb

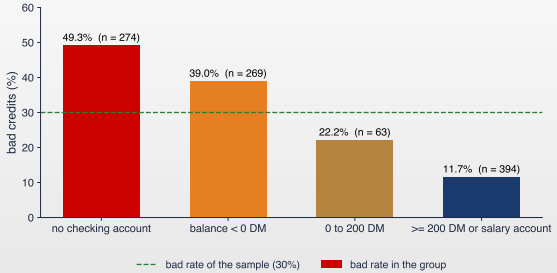
Datele South German Credit

- 1 000 de credite de consum ale unei mari bănci regionale din sudul Germaniei, 1973–1975 South German Credit (2020); Grömping (2019)
  - rezultatul: contractul a fost respectat (bun-platnic) sau nu (rău-platnic, **nerambursare** $= 1$): 300 de credite neperformante și 700 performante
  - 20 de variabile: contul curent, durata, istoricul de credit, destinația, suma (DM, mărci germane), economiile, vechimea în muncă, vîrsta și altele
- Un **eșantion stratificat**: creditele neperformante sînt puternic suprareprezentate; rata lor în bancă era de aproximativ 5% Grömping (2019)
  - toți debitorii trecuseră de verificările băncii: solicitanții respinși lipsesc (reject inference, secțiunea 8)
  - costul acceptării unui credit neperformant este stabilit la de cinci ori costul respingerii unuia performant
- Licența CC BY 4.0; salvate o singură dată în repository-ul cursului (data/credit)

Erori de codificare în datele German Credit

- Datele au fost folosite zeci de ani în versiunea UCI „Statlog (German Credit)", cu etichete greșite ale codurilor Grömping (2019)
  - în acea versiune, „fără cont curent" era grupul cel mai sigur: 394 de credite, 11,7% neperformante
  - în datele corectate, acele 394 de credite au un sold de cel puțin 200 DM sau un cont de salariu; „fără cont curent" este grupul cel mai riscant (49,3% neperformante)
- Cifrele erau corecte, etichetele greșite: fiecare model funcționa, fiecare interpretare a acestei variabile era inversată
  - verificări: comparăm tabelele de frecvențe cu sursa originală (aici cartea lui Fahrmeir și Hamerle) și ne întrebăm dacă semnele au sens economic
- Folosim în tot capitolul datele corectate South German Credit

Rata creditelor neperformante după contul curent



- Ponderea creditelor neperformante în fiecare grup; linia întreruptă: media eșantionului, 30%
- Interpretare: riscul scade de la 49,3% (fără cont) la 11,7% (sold de cel puțin 200 DM): o singură variabilă separă deja bine solicitanții

SFM_ch12_woe_iv

Dezechilibrul claselor și suprareprezentarea

- Nerambursările sînt rare: cîteva procente din creditele unei bănci; un eșantion cu puțini rău-platnici spune puțin despre ei
  - de aceea băncile suprareprezintă rău-platnicii (aici 30% în loc de circa 5%)
- **Paradoxul acurateței**: regula „toți sînt bun-platnici" are dreptate pentru 70% din acest eșantion și pentru 77,9% din datele din Taiwan
  - deci acuratețea singură este o măsură slabă; avem nevoie de măsuri care privesc separat rău-platnicii și bun-platnicii (secțiunea 7)
- **Datele cardurilor de credit din Taiwan** Default of Credit Card Clients (2009); Yeh & Lien (2009): 30 000 de deținători de card, nerambursarea plății din octombrie 2005, 22,1% nerambursări
  - variabile: limita de credit, vîrsta, starea plăților din aprilie–septembrie 2005, facturile și plățile; le folosim pentru calibrare și echitate

Eșantionul de estimare și eșantionul de test

- **Împărțire stratificată**: 700 de credite pentru estimare, 300 pentru test, cu 30% rău-platnici în ambele
  - orice alegere (intervale, WoE, variabile, coeficienți) se face doar pe eșantionul de estimare
  - eșantionul de test se folosește o singură dată, la final: joacă rolul viitorilor solicitanți
- **Leakage** (scurgerea de informație): informație din eșantionul de test care intră în model, de exemplu intervale alese pe toate datele
  - face rezultatul testului prea optimist; cu 300 de credite de test, o singură împărțire este și zgomotoasă, așa că adăugăm validarea încrucișată (secțiunea 7)

Weight of Evidence (WoE)

- Împărțim o variabilă în intervale (clase) $i = 1, \ldots, m$; g_i, b_i: bun-platnicii și rău-platnicii din intervalul i; G, B: totalurile lor
  - $\mathrm{WoE}_i = \ln \dfrac{g_i/G}{b_i/B}$: logaritmul raportului dintre ponderea bun-platnicilor și ponderea rău-platnicilor în interval
- Sensul: $\mathrm{WoE}_i = \ln \dfrac{g_i}{b_i} - \ln \dfrac{G}{B}$: logaritmul șansei de a fi bun-platnic în interval minus cel din întregul eșantion
  - $\mathrm{WoE} > 0$: mai sigur decît media; $\mathrm{WoE} < 0$: mai riscant; $\mathrm{WoE} = 0$: nicio informație
  - un interval fără bun-platnici sau fără rău-platnici ar da $\pm\infty$: adăugăm 0,5 la ambele frecvențe
- **Information Value**: $\mathrm{IV} = \sum_i \left(\dfrac{g_i}{G} - \dfrac{b_i}{B} \right) \mathrm{WoE}_i \geq 0$
  - o divergență Kullback–Leibler simetrică între distribuțiile bun-platnicilor și ale rău-platnicilor pe intervale
  - regula practică Siddiqi (2006): sub 0,02 inutilă, 0,02–0,1 slabă, 0,1–0,3 medie, peste 0,3 puternică

Exemplu rezolvat: WoE și IV pentru contul curent

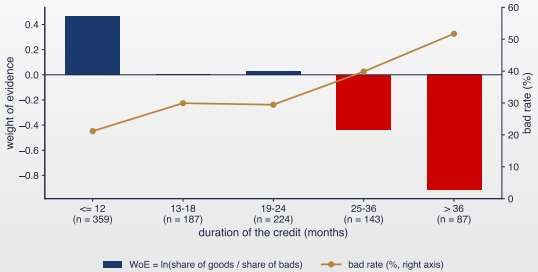
Contul curent	bun-pl.	rău-pl.	g_i/G	b_i/B	WoE	termen IV
fără cont curent	139	135	0,199	0,450	−0,818	0,206
sold $<$ 0 DM	164	105	0,234	0,350	−0,401	0,046
0–200 DM	49	14	0,070	0,047	0,405	0,009
$\geq$ 200 DM sau salariu	348	46	0,497	0,153	1,176	0,404
total	700	300	1	1		0,666

- **Pas cu pas**, ultimul rînd: $g/G = 348/700 = 0{,}497$, $b/B = 46/300 = 0{,}153$
  - WoE $= \ln(0{,}497/0{,}153) = 1{,}176$; termenul IV $= (0{,}497 - 0{,}153) \times 1{,}176 = 0{,}404$
- Interpretare: IV $= 0{,}666 > 0{,}3$: o variabilă puternică; WoE scade monoton de la grupul cel mai sigur la cel mai riscant

Discretizarea (binning): reguli practice

- Rolul intervalelor: surprind efectele neliniare, limitează influența valorilor extreme și dau fiecărui solicitant un tabel mic de puncte
  - variabile numerice: cîteva intervale (aici durata ≤ 12, 13–18, 19–24, 25–36, > 36 de luni); categoriale: categoriile
- Reguli Siddiqi (2006); Thomas, Crook & Edelman (2017)
  - cel puțin aproximativ 5% din eșantion în fiecare interval; unim categoriile mici
  - WoE trebuie să varieze într-un mod cu sens economic (adesea monoton)
  - intervalele și WoE se estimează pe eșantionul de estimare și se aplică neschimbate noilor solicitanți
- O variabilă WoE intră în modelul logit cu un singur coeficient: scorecard-ul rămîne aditiv și ușor de citit

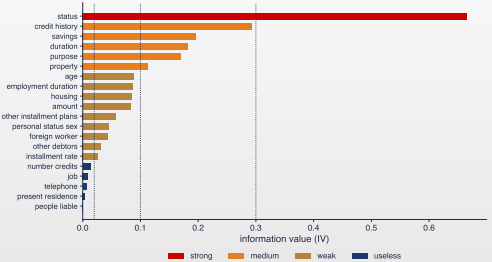
WoE pentru durata creditului



- Eșantionul complet; bare: WoE pentru fiecare interval de durată; linia: rata creditelor neperformante (axa din dreapta); IV = 0,182 (medie)
- Interpretare: creditele de pînă la un an sînt mai sigure (WoE = 0,47, 21% neperformante); peste trei ani, rata ajunge la 52% (WoE = −0,92); 13–24 de luni sînt în medie

SFM_ch12_woe_iv

Information Value pentru cele 20 de variabile



- Eșantionul complet; linii punctate: pragurile 0,02; 0,1 și 0,3; pe eșantionul de estimare, 6 variabile au IV $\geq$ 0,1 și intră în scorecard
- Interpretare: contul curent (IV = 0,666) și istoricul de credit (0,293) domină; telefonul, vechimea la adresă și persoanele în întreținere nu aduc informație
- IV privește o singură variabilă: două variabile puternice, dar corelate, nu își adună informația

SFM_ch12_woe_iv

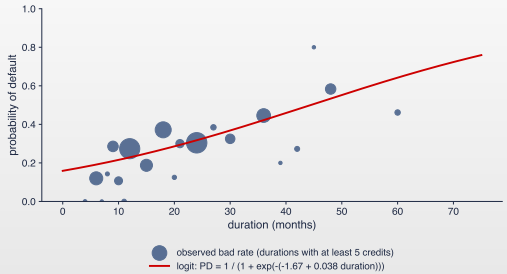
Șansa (odds) și logaritmul șansei

- Un model liniar $P(Y = 1) = \beta_0 + x^\top \beta$ poate da probabilități sub 0 sau peste 1
  - modelăm o transformare a lui p care poate lua orice valoare reală
- **Șansa** (odds) de nerambursare: $\frac{p}{1-p} \in (0, \infty)$; **logaritmul șansei** (logit): $\ln \frac{p}{1-p} \in (-\infty, \infty)$
  - $p = 0{,}20$: șansa $= 0{,}2/0{,}8 = 0{,}25$ (o nerambursare la patru credite rambursate); logaritmul șansei $= \ln 0{,}25 = -1{,}386$
  - $p = 0{,}5$: șansa 1, logaritmul șansei 0; logit-ul este simetric: $\mathrm{logit}(1-p) = -\mathrm{logit}(p)$
- Bancherii vorbesc adesea despre **șansa bun:rău** $(1-p)/p$: aici 4:1

Modelul logit

- $\ln \dfrac{p_i}{1-p_i} = \beta_0 + x_i^\top \beta = \eta_i, \quad p_i = P(Y_i = 1 \mid x_i) = \dfrac{1}{1+e^{-\eta_i}}$ Cox (1958)
  - $x_i = (x_{i1}, \ldots, x_{ik})^\top$: vectorul celor k caracteristici ale solicitantului i; $\beta = (\beta_1, \ldots, \beta_k)^\top$: vectorul coeficienților lor, același pentru toți solicitanții
  - β_0: termenul liber; $\eta_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik}$: predictorul liniar, logaritmul șansei solicitantului i
- **Verosimilitatea maximă** (Capitolul 6): $\ell(\beta) = \sum_i [y_i \ln p_i + (1-y_i)\ln(1-p_i)]$
  - condiția de ordinul întîi $\sum_i (y_i - p_i) x_i = 0$: fără formulă explicită, se rezolvă prin Newton–Raphson
  - erorile standard din inversa lui $\sum_i p_i(1-p_i) x_i x_i^\top$; testul Wald $z = \hat\beta_j / \mathrm{se}(\hat\beta_j)$
- Modelul logit este folosit și pentru falimentul companiilor, începînd cu Ohlson (1980); modelul probit folosește $\Phi(\eta)$ în locul funcției logistice
  - logit și probit dau aproape aceleași PD; logit-ul are interpretarea prin șanse

O curbă logistică: PD și durata



- Logit doar cu durata: $\hat{\beta}_{\mathbf{0}} = -1{,}666$, $\hat{\beta}_{\mathbf{1}} = 0{,}0375$ pe lună (se 0,0057); punctele: ratele observate
- Interpretare: fiecare an în plus înmulțește șansa de nerambursare cu $e^{\mathbf{12}\times\mathbf{0{,}0375}} = 1{,}57$; PD crește de la 22,9% la 12 luni la 53,4% la 48 de luni

SFM_ch12_logit_lda

Interpretarea coeficienților

- β_j: modificarea **logaritmului șansei** cînd x_j crește cu o unitate, celelalte variabile fiind fixe
  - e^{β_j}: **raportul șanselor** (odds ratio): șansa se înmulțește cu e^{β_j}; $100(e^{\beta_j} - 1)\%$ este modificarea șansei
  - un interval de încredere de 95%: $\exp(\hat{\beta}_j \pm 1{,}96\,\mathrm{se})$
- Efectul asupra **probabilității** depinde de poziția solicitantului
  - $\partial p / \partial x_j = p(1-p)\beta_j$: maxim la $p = 0{,}5$, mic pentru solicitanții foarte siguri sau foarte riscanți
- O variabilă indicator (0/1): e^{β_j} compară șansa grupului cu cea a grupului de referință
  - interpretarea cere un model care conține variabilele relevante: o variabilă omisă schimbă coeficienții

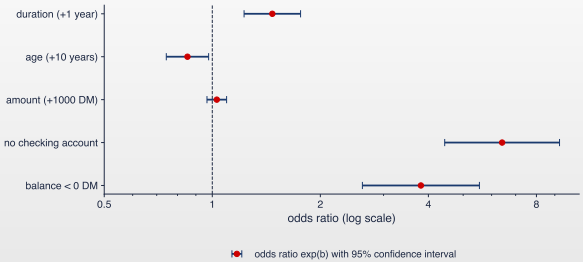
Un model logit mic pe eșantionul complet

Variabila	$\hat{\beta}$	se	z	p	raportul șanselor (CI 95%)
termenul liber	−2,101	0,311	−6,75	$<0,001$	
durata (ani)	0,385	0,092	4,17	$<0,001$	1,47 (1,23–1,76)
vîrsta (decenii)	−0,159	0,069	−2,30	0,022	0,85 (0,74–0,98)
suma (1000 DM)	0,029	0,032	0,90	0,368	1,03 (0,97–1,10)
fără cont curent	1,859	0,188	9,89	$<0,001$	6,42 (4,44–9,28)
sold $<$ 0 DM	1,338	0,192	6,99	$<0,001$	3,81 (2,62–5,55)

- $n = 1\,000$; log-verosimilitatea $-526{,}1$ față de $-610{,}9$ doar cu termenul liber: LR $= 169{,}6$ cu 5 grade de libertate
- Grupul de referință al indicatorilor de cont: un sold nenegativ (0–200 DM, sau cel puțin 200 DM ori cont de salariu)

SFM_ch12_logit_lda

Rapoartele șanselor cu intervale de încredere



- Fiecare punct: $e^{\hat{\beta}}$; bare: intervale de 95%; scală logaritmică; linia verticală: niciun efect (1)
- Interpretare: fără cont curent, șansa de nerambursare este de 6,42 ori cea a grupului de referință; un sold negativ o înmulțește cu 3,81
- Un an în plus de durată crește șansa cu 47%; zece ani în plus de vîrstă o reduc cu aproximativ 15%
- Suma nu este semnificativă odată ce durata este în model (cele două sînt corelate)

SFM_ch12_logit_lda

Exemplu rezolvat: PD a unui solicitant

- Solicitantul: 36 de luni, 30 de ani, 4000 DM, fără cont curent; coeficienții din tabelul anterior, rotunjiți
  - $\eta = -2{,}101 + 0{,}385 \times 3 + (-0{,}159) \times 3 + 0{,}029 \times 4 + 1{,}859$
  - $\eta = -2{,}101 + 1{,}155 + (-0{,}477) + 0{,}116 + 1{,}859 = 0{,}552$
- PD $= 1/(1 + e^{-0{,}552}) = 63{,}5\%$; șansa $e^{0{,}552} = 1{,}74$
  - același solicitant cu un sold nenegativ: $\eta = -1{,}307$, PD $= 21{,}3\%$
  - un an în plus de durată: $\Delta p \approx p(1-p)\beta \approx 8{,}9$ puncte procentuale
- Atenție: această PD se referă la un eșantion cu 30% rău-platnici, nu la populația băncii (corecția pentru proporția a priori, peste două slide-uri)

Scorecard-ul logit pe variabile WoE

Variabila WoE	IV (estimare)	$\hat{\beta}$	se
termenul liber		−0,850	0,095
contul curent	0,591	−0,829	0,127
istoricul de credit	0,215	−0,675	0,206
destinația creditului	0,205	−0,937	0,217
economiile	0,166	−0,726	0,246
durata	0,159	−1,082	0,239
vechimea la angajator	0,122	−0,873	0,265

- Eșantionul de estimare (700 de credite); fiecare variabilă este înlocuită cu WoE al intervalului ei; logit pentru nerambursare
  - toți coeficienții sînt negativi: un WoE mai mare (interval mai sigur) reduce logaritmul șansei de nerambursare
  - un coeficient de -1 ar însemna că variabila adaugă exact propriul WoE; valorile între $-0{,}7$ și $-1{,}1$ reflectă suprapunerea dintre variabile
- Eșantionul de test: AUC $= 0{,}801$ (secțiunea 7); log-verosimilitatea în eșantion $-347{,}2$

SFM_ch12_scorecard

Corecția pentru suprareprezentarea rău-platnicilor

- Dacă rău-platnicii și bun-platnicii sînt eșantionați cu rate diferite, se schimbă doar termenul liber al modelului logit
  - $\ln \text{șansa}_{\text{pop}} = \ln \text{șansa}_{\text{eșantion}} + \ln \dfrac{\pi/(1-\pi)}{\pi_s/(1-\pi_s)}$; π: rata din populație, π_s: rata din eșantion
  - pantele rămîn neschimbate: ordonarea solicitanților nu depinde de eșantionare
- **Pas cu pas**: $\pi_s = 0{,}30$, $\pi = 0{,}05$: corecția $= \ln(0{,}05/0{,}95) - \ln(0{,}30/0{,}70) = -2{,}944 - (-0{,}847) = -2{,}097$
  - o PD de 30% în eșantion devine 5%; o PD de 50% devine 10,9%
  - solicitantul din exemplul rezolvat: $0{,}552 - 2{,}097 = -1{,}545$, PD $= 17{,}6\%$ în populația băncii, în loc de 63,5%
- Interpretare: AUC și Gini nu sînt afectate de corecție; EL și capitalul sînt afectate

Funcția discriminantă a lui Fisher (1936)

- Fisher (1936): găsim combinația liniară $w^\top x$ care separă cel mai bine două grupuri
  - a folosit măsurători ale florilor de iris; Durand a aplicat metoda creditelor cinci ani mai tîrziu
- Criteriul Fisher: $J(w) = \frac{[w^\top (m_1 - m_0)]^2}{w^\top S_W w}$
  - m_0, m_1: vectorii mediilor pentru bun-platnici și rău-platnici; S_W: matricea de covariață comună din interiorul grupurilor
  - maximul: $w \propto S_W^{-1}(m_1 - m_0)$ Härdle & Simar (2019), cap. 14
- Criteriul nu cere nicio distribuție: este o proiecție care maximizează distanța dintre grupuri raportată la dispersia lor



R.A. Fisher, 1913

Foto: autor necunoscut (1913); domeniu public; Wikimedia Commons



Iris versicolor

Foto: Ermell (2020); CC BY-SA 4.0; Wikimedia Commons

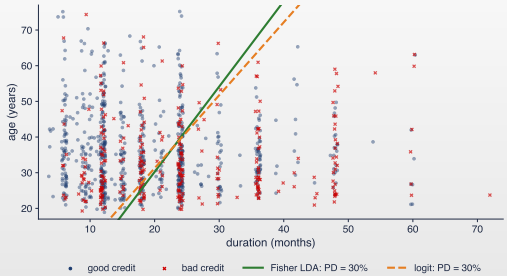
LDA ca model de probabilitate

- Presupunem $x \mid Y = k \sim N(m_k, \Sigma)$, aceeași Σ în ambele grupuri, și probabilitatea a priori $\pi_1 = P(Y = 1)$
  - Bayes: $\ln \dfrac{P(Y=1 \mid x)}{P(Y=0 \mid x)} = \underbrace{\ln \dfrac{\pi_1}{1-\pi_1} - \frac{1}{2}(m_1 + m_0)^\top w}_{c} + w^\top x, \quad w = \Sigma^{-1}(m_1 - m_0)$
  - logaritmul șansei este **liniar în** x, exact ca în modelul logit; probabilitatea a priori intră doar în termenul liber
- Diferența este în estimare
  - LDA: folosim direct mediile de selecție și covarianța comună (formulă explicită)
  - logit: maximizăm verosimilitatea condiționată a lui Y dată fiind x; nicio ipoteză asupra distribuției lui x
- Regula de clasificare: clasificăm solicitantul ca „rău-platnic" dacă $c + w^\top x > \ln(C_{FP}/C_{FN})$ (costuri, secțiunea 7)

Exemplu rezolvat: direcția Fisher cu două variabile

- x = (durata în luni, vîrsta în ani), eșantionul complet: 700 de bun-platnici, 300 de rău-platnici
  - $m_{\mathbf{0}} = (19{,}21; 36{,}22)$, $m_{\mathbf{1}} = (24{,}86; 33{,}96)$, $m_{\mathbf{1}} - m_{\mathbf{0}} = (5{,}65; -2{,}26)$
  - $S_W = \begin{pmatrix} 138{,}8 & -2{,}46 \\ -2{,}46 & 127{,}9 \end{pmatrix}$, $\det S_W = 17756$
- **Pas cu pas**: $S_W^{-\mathbf{1}} = \dfrac{1}{\det S_W} \begin{pmatrix} 127{,}9 & -(-2{,}46) \\ -(-2{,}46) & 138{,}8 \end{pmatrix}$, $w = S_W^{-\mathbf{1}}(m_{\mathbf{1}} - m_{\mathbf{0}}) = (0{,}0404; -0{,}0169)$
  - logit cu aceleași două variabile: pantele $(0{,}0375; -0{,}0183)$
  - raportul durată/vîrstă: LDA $-2{,}39$, logit $-2{,}05$: o lună de durată cîntărește cît aproximativ doi ani de vîrstă, în sens invers

Frontiera Fisher și frontiera logit



- Durata și vîrsta fiecărui credit (cu o mică perturbare); liniile: punctele cu PD = 30% (rata din eșantion) pentru LDA și pentru logit
- Interpretare: ambele frontiere sînt drepte cu aproape aceeași pantă; cei doi nori se suprapun mult, deci două variabile sînt departe de a fi suficiente

SFM_ch12_logit_lda

Compararea modelelor logit și LDA

	Logit	LDA (Fisher)
modelul	doar $P(Y \mid x)$	$P(x \mid Y)$ Normală, Σ comun
estimarea	verosimilitate maximă, iterativ	medii și covariață, formulă explicită
dacă x este Normal	consistent, mai puțin eficient	mai eficient
indicatori, WoE, x asimetric	potrivit	ordonarea bună, PD pot fi greșite
AUC de test (variabile WoE)	0,801	0,801

- Eșantionul de test: corelația dintre cele două scoruri (logaritmul șansei) este 0,999; testul DeLong de egalitate a AUC: $z = 0{,}18$, $\mathrm{p} = 0{,}86$ DeLong et al. (1988)
- Interpretare: pentru ordonare, cele două metode sînt practic identice; logit-ul este preferat pentru PD, deoarece nu cere variabile Normale

SFM_ch12_logit_lda

Scalarea unui scorecard

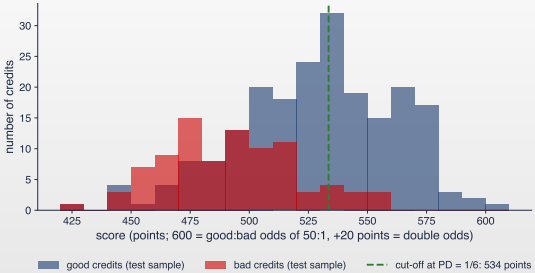
- **Scorecard** (grila de punctaj): un tabel care dă puncte fiecărui atribut; scorul unui solicitant este suma punctelor Siddiqi (2006)
  - scor mai mare = risc mai mic; scorul este o funcție liniară de logaritmul șansei bun:rău
- $\text{Scor} = \text{Offset} + \text{Factor} \times \ln \frac{1-p}{p}$, două convenții fixează scala
  - **PDO** (points to double the odds): +PDO puncte $\Leftrightarrow$ șansa bun:rău se dublează: $\text{Factor} = \text{PDO}/\ln 2$
  - un scor de bază la o șansă dată: $\text{Offset} = \text{Baza} - \text{Factor} \ln(\text{șansa}_0)$
- Aici: PDO = 20, 600 de puncte la șansa 50:1
  - $\text{Factor} = 20/\ln 2 = 28{,}85$; $\text{Offset} = 600 - 28{,}85 \times \ln 50 = 600 - 28{,}85 \times 3{,}912 = 487{,}12$

Exemplu rezolvat: scoruri și puncte

- **De la PD la scor**: PD $= 5\%$, șansa 19:1: Scorul $= 487{,}12 + 28{,}85 \times \ln 19 = 487{,}12 + 28{,}85 \times 2{,}944 = 572{,}1$
  - PD $= 30\%$ (șansa 7:3): $487{,}12 + 28{,}85 \times 0{,}847 = 511{,}6$
- **De la scor la PD**: 560 de puncte: ln șansa $= (560 - 487{,}12)/28{,}85 = 2{,}526$, șansa 12,5:1, PD $= 7{,}4\%$
- **Punctele unui atribut** ($k = 6$ variabile; logit pentru nerambursare cu termenul liber β_0 și panta β_j)
  - $\text{puncte}_{ij} = -(\beta_j \, \text{WoE}_{ij} + \beta_0/k)\,\text{Factor} + \text{Offset}/k$; semnul minus transformă logaritmul șansei de nerambursare în logaritmul șansei bun:rău
  - contul curent (WoE din estimare): fără cont 66 de puncte, sold < 0 77, 0–200 DM 97, cel puțin 200 DM sau salariu 110
- Corecția a priori în puncte: pentru o populație cu 5% rău-platnici, logaritmul șansei de nerambursare scade cu 2,097, deci logaritmul șansei bun:rău crește cu 2,097 și fiecare scor crește cu $28{,}85 \times 2{,}097 = 60{,}5$ puncte

SFM_ch12_scorecard

Scorurile eșantionului de test



- 300 de credite de test; scoruri de la 422 la 607 puncte; media: bun-platnici 529, rău-platnici 492; linia întreruptă: pragul la PD $= 1/6$ (secțiunea 7), 534 de puncte
- Interpretare: cele două distribuții sînt deplasate, dar se suprapun: orice prag respinge unii bun-platnici și acceptă unii rău-platnici; măsurile de validare cuantifică această suprapunere

SFM_ch12_scorecard

Matricea de confuzie

	prezis rău (respins)	prezis bun (acceptat)
rău-platnic real	TP	FN
bun-platnic real	FP	TN

- Un prag c: respingem dacă PD $> c$
- TP (true positive): un rău-platnic respins; FN (false negative): un rău-platnic acceptat
- FP (false positive): un bun-platnic respins; TN (true negative): un bun-platnic acceptat

- Ratele
  - **TPR** (true positive rate, sensibilitatea) $= \mathrm{TP}/(\mathrm{TP}+\mathrm{FN})$: ponderea rău-platnicilor identificați
  - **FPR** (false positive rate) $= \mathrm{FP}/(\mathrm{FP}+\mathrm{TN}) = 1-$ specificitatea: ponderea bun-platnicilor respinși
  - precizia $= \mathrm{TP}/(\mathrm{TP}+\mathrm{FP})$; acuratețea $= (\mathrm{TP}+\mathrm{TN})/n$
- Fiecare rată depinde de c; măsurile următoare rezumă toate pragurile deodată

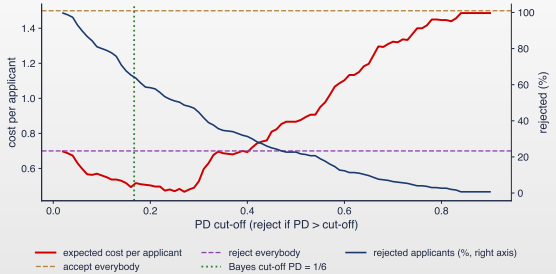
Două praguri pe eșantionul de test

prag	TP	FN	FP	TN	acurateţe	TPR	FPR	cost
PD $> 0{,}5$	43	47	25	185	76,0%	47,8%	11,9%	0,867
PD $> 1/6$	82	8	112	98	60,0%	91,1%	53,3%	0,507

- Costul pe solicitant $= (5\,\mathrm{FN} + 1\,\mathrm{FP})/n$: acceptarea unui rău-platnic costă de cinci ori mai mult decît respingerea unui bun-platnic Grömping (2019)
  - pragul Bayes minimizează costul așteptat: respingem dacă $p\,C_{FN} > (1-p)\,C_{FP}$, adică $p > C_{FP}/(C_{FP} + C_{FN}) = 1/6$
- Interpretare: pragul de 1/6 are o acuratețe mai mică (60,0% față de 76,0%), dar un cost mult mai mic (0,507 față de 0,867): identifică 91,1% din rău-platnici

SFM_ch12_validation_irb

Costul așteptat și alegerea pragului



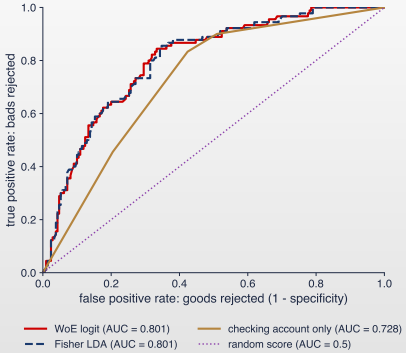
- Eșantionul de test; roșu: costul pe solicitant; linii întrerupte: acceptăm pe toată lumea (1,50) sau respingem pe toată lumea (0,70); albastru: ponderea respinșilor (axa din dreapta)
- Interpretare: costul este aproape constant între PD 0,1 și 0,3; minimul pe aceste 300 de credite este la 0,27, aproape de pragul Bayes de 1/6, care se alege înainte de a vedea datele de test

SFM_ch12_validation_irb

Curba ROC și AUC

- **ROC** (receiver operating characteristic): punctele (FPR(c), TPR(c)) pentru toate pragurile c
  - de la (0, 0) (nu respingem pe nimeni) la (1, 1) (respingem pe toată lumea); diagonala este un scor aleator
- **AUC** (aria de sub curbă) $= P(s_{\text{rău}} > s_{\text{bun}}) + \frac{1}{2}P(s_{\text{rău}} = s_{\text{bun}})$ Hanley & McNeil (1982)
  - s: scorul de risc (aici PD); se calculează numărînd perechile: ponderea perechilor (rău, bun) ordonate corect
  - statistica Mann–Whitney împărțită la $n_1 n_0$; 0,5: nicio discriminare, 1: perfectă
- **Gini** $= 2\,\mathrm{AUC} - 1$; AUC nu depinde de rata rău-platnicilor sau de o transformare monotonă a scorului
  - eroarea standard și testele pentru AUC: DeLong et al. (1988)

Curbele ROC pe eșantionul de test



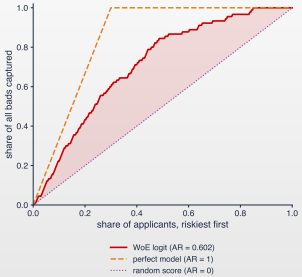
- 300 de credite de test, 18 900 de perechi (rău, bun)
- Logit WoE: AUC $= 0{,}801$, CI 95% $[0{,}749; 0{,}854]$ (DeLong, se $0{,}027$); Gini $= 0{,}603$
- Doar contul curent: AUC $= 0{,}728$; testul DeLong față de scorecard: $z = 2{,}86$, $p = 0{,}004$
- Interpretare: un rău-platnic ales la întîmplare are o PD mai mare decît un bun-platnic ales la întîmplare în aproximativ 80% din perechi; celelalte cinci variabile aduc o îmbunătățire semnificativă față de contul curent

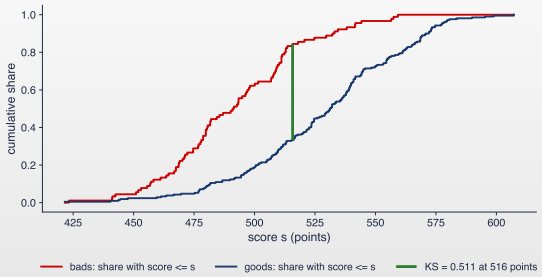
SFM_ch12_validation_irb

Curba CAP, raportul de acuratețe și KS

- **CAP** (cumulative accuracy profile): ordonăm solicitanții de la cel mai riscant; reprezentăm ponderea tuturor rău-platnicilor printre primii $x\%$
  - modelul perfect: toți rău-platnicii primii, curba atinge 1 la $x =$ rata rău-platnicilor; aleator: diagonala
  - **AR** (accuracy ratio, raportul de acuratețe) $= \dfrac{\text{aria dintre CAP-ul modelului și diagonală}}{\text{aria dintre CAP-ul perfect și diagonală}}$; pentru un scor continuu AR $=$ Gini
- **KS** (Kolmogorov–Smirnov): $\max_s |F_{\text{rău}}(s) - F_{\text{bun}}(s)|$, cea mai mare distanță dintre distribuțiile scorurilor rău-platnicilor și bun-platnicilor
  - $F_{\text{rău}}$, F_{bun}: funcțiile de repartiție empirice ale scorurilor în cele două grupuri; KS $\in [0, 1]$
  - citit pe curba ROC: KS $= \max_c (\text{TPR} - \text{FPR})$; scorul la care apare este un prim prag natural
- Toate cele trei măsoară **discriminarea** (ordonarea), nu calibrarea

CAP și KS pe eșantionul de test





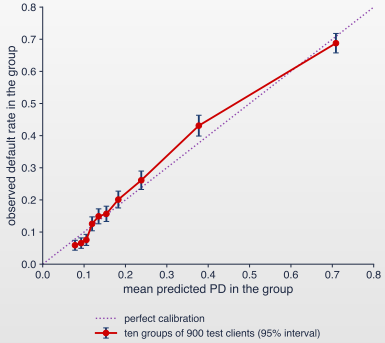
- AR = 0,602 (Gini = 0,603: același, cu excepția rotunjirii); KS = 0,511 la 516 puncte
- Interpretare: respingînd cei mai riscanți 30%, identificăm 60% din rău-platnici (54 din 90); la 516 puncte, scorecard-ul separă cel mai bine rău-platnicii de bun-platnici

SFM_ch12_validation_irb

Calibrarea și scorul Brier

- **Calibrarea**: dintre solicitanții cu PD $\approx p$, aproximativ o pondere p intră în nerambursare
  - diagrama de calibrare: ordonăm după PD, formăm grupuri (de exemplu zece), reprezentăm PD medie față de rata observată
  - Hosmer–Lemeshow Hosmer & Lemeshow (1980): $\sum_g \frac{(O_g - n_g \bar{p}_g)^2}{n_g \bar{p}_g (1 - \bar{p}_g)} \approx \chi^2(G-2)$
  - G grupuri; n_g: mărimea, O_g: nerambursările observate, $\bar{p}_g$: PD medie a grupului g; o valoare mare: calibrare slabă
- **Scorul Brier** Brier (1950): $\mathrm{BS} = \frac{1}{n}\sum_i (p_i - y_i)^2$, eroarea pătratică medie a PD; o valoare mai mică este mai bună
  - referința: o PD constantă egală cu rata din eșantionul de estimare; cîștigul $= 1 - \mathrm{BS}/\mathrm{BS}_{\mathbf{ref}}$
  - eșantionul de test al scorecard-ului: BS = 0,162 față de 0,210 (cîștig 23%); HL cu 10 grupuri = 6,520, p = 0,589
- Scorul Brier recompensează atît ordonarea, cît și calibrarea; AUC măsoară doar ordonarea

Calibrarea pe datele din Taiwan



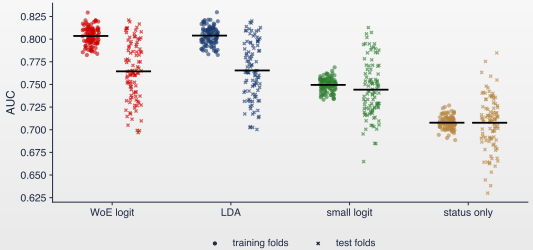
- Logit pe 30 000 de deținători de card (70% estimare); eșantion de test de 9 000; zece grupuri de cîte 900 după PD; bare: intervale de 95%
- AUC de test $= 0{,}777$ (estimare 0,774); Brier $= 0{,}136$ față de 0,172 (cîștig 21%)
- Hosmer–Lemeshow $= 40{,}660$, $p < 0{,}001$: respins, deși punctele sînt aproape de diagonală
- Interpretare: cu 9000 de clienți, testul detectează abateri mici (grupul cel mai sigur: PD 7,8%, observat 5,9%); judecăm mărimea abaterilor, nu doar p-value-ul

SFM_ch12_validation_irb

Validarea în afara eșantionului și validarea încrucișată

- Performanța **în eșantion** este optimistă: modelul s-a adaptat la zgomotul datelor de estimare (supraajustare)
  - un eșantion de test de 300 de credite dă AUC $= 0{,}801$ cu o eroare standard de $0{,}027$: o singură împărțire este zgomotoasă
- **Validarea încrucișată în k părți**: împărțim în $k = 5$ părți; estimăm pe patru, testăm pe a cincea; rotim; repetăm de 20 de ori
  - întreaga procedură se repetă în fiecare parte: intervale, WoE, selecția după IV și coeficienții
  - raportăm media și dispersia celor 100 de AUC de test, precum și diferența dintre AUC de estimare și de test
- Validarea **out-of-time**: estimăm pe cereri mai vechi, testăm pe cereri ulterioare (peste două slide-uri)

AUC de estimare și de test în validarea încrucișată repetată



- 20 de repetări a cîte 5 părți; puncte: părțile de estimare, cruci: părțile de test; liniile negre: mediile
- Logit WoE: estimare 0,804, test 0,764 (sd 0,032); LDA: 0,804, 0,765; logit mic: 0,749, 0,744; contul curent: 0,708
- Interpretare: scorecard-ul pierde 0,039 din AUC în afara eșantionului, logit-ul cu cinci variabile aproape nimic; împărțirea unică de mai sus a fost norocoasă (0,801 față de 0,764 în medie)

SFM_ch12_validation_irb

Validarea out-of-time și deriva populației

- Un scorecard este folosit ani de zile pe solicitanți care vin **după** perioada de estimare
  - economia, produsele și solicitanții se schimbă: relația dintre x și nerambursare se modifică în timp
  - testul out-of-time: estimăm, de exemplu, pe 2019–2021, testăm pe 2022–2023; comparăm AUC și calibrarea de la un an la altul
- Monitorizarea: distribuția scorurilor și a fiecărei variabile se compară cu perioada de estimare
  - o deplasare cu ordonare stabilă cere recalibrare (termenul liber); o pierdere a ordonării cere un model nou
- Limitarea datelor: niciunul dintre seturile de date ale capitolului nu conține data cererilor, deci un test out-of-time nu este posibil aici; putem folosi doar împărțiri aleatoare
- Lectură suplimentară despre o validare out-of-time, pentru activele cripto „zombie”: Będowska-Sójka, Wójcik & Pele (2026)

Studiu de caz: compararea metodelor de scoring

- Lessmann et al. (2015): 41 de clasificatori, șase măsuri de acuratețe, opt seturi de date de retail (printre ele datele germane de credit din UCI)
  - o actualizare, după zece ani, a comparației lui Baesens et al. (2003)
- Rezultatele
  - mai mulți clasificatori prezic semnificativ mai bine decît regresia logistică, standardul industriei; ansamblurile eterogene sînt cele mai bune
  - scorecard-urile mai precise pot aduce cîștiguri financiare importante
  - măsurile uzuale de acuratețe dau semnale asemănătoare; random forests sînt recomandate ca reper pentru orice metodă nouă
- A învinge doar logit-ul nu mai este o dovadă de progres; pe cele 1000 de credite, logit și LDA diferă cu mai puțin de o eroare standard a AUC; Capitolul 13 adaugă arbori și rețele neuronale

Reject inference

- Rezultatul este cunoscut doar pentru solicitanții **acceptați**; modelul se folosește pe **toți** solicitanții („through the door")
  - un eșantion selectat de scorul vechi: lipsesc solicitanții cei mai riscanți și rezultatele lor
  - debitorii din South German Credit trecuseră toți de verificările băncii Grömping (2019)
- **Reject inference**: metode care încearcă să corecteze lipsa rezultatelor Banasik, Crook & Thomas (2003); Thomas, Crook & Edelman (2017)
  - reponderarea: dăm o pondere mai mare acceptaților care seamănă cu respinșii
  - augmentarea și parcelarea: atribuim rezultate respinșilor pe baza unui model
  - sursa cea mai fiabilă: un mic eșantion aleator de solicitanți acceptați indiferent de scor

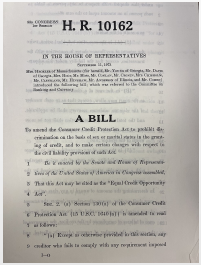
Un experiment de reject inference pe datele din Taiwan

- Un „scor vechi" (întîrzierea din septembrie și limita) acceptă cei mai buni 70% dintre clienții de estimare (15 085); rata lor de nerambursare este 13,2%, a respinșilor 44,9%
  - printre acceptați nimeni nu a întîrziat în septembrie: variabila cea mai puternică nu poate fi estimată și iese din model
- Eșantionul de test (toți solicitanții): AUC al modelului estimat doar pe acceptați 0,736, al celui estimat pe toți 0,777, al scorului vechi 0,743
  - clienții de test respinși: rata observată 46,1%; PD medie a modelului estimat pe acceptați 29,6%, a modelului complet 44,1%
- Interpretare: estimat doar pe acceptați, noul model ordonează mai slab decît scorul vechi și subestimează riscul respinșilor cu aproximativ 36%

SFM_ch12_reject_fairness

Echitatea (fairness) în scoringul de credit

- Legea americană interzice discriminarea în creditare după sex, stare civilă, rasă, religie, origine națională sau vîrstă Equal Credit Opportunity Act (FTC)
  - UE: scoringul de credit al persoanelor este o utilizare cu risc ridicat a AI (documentare, supraveghere umană) Comisia Europeană, AI Act
- Excluderea variabilei protejate nu este suficientă: alte variabile o pot înlocui
  - **paritatea demografică**: rate egale de aprobare între grupuri
  - **egalitatea de șanse**: rate egale de aprobare pentru bun-platnici Hardt, Price & Srebro (2016)
  - **calibrarea în fiecare grup**: aceeași PD înseamnă același risc în fiecare grup
- Cînd ratele de nerambursare diferă între grupuri, aceste criterii nu pot fi toate îndeplinite simultan: alegerea ține de politica băncii și a autorităților



H. R. 10162

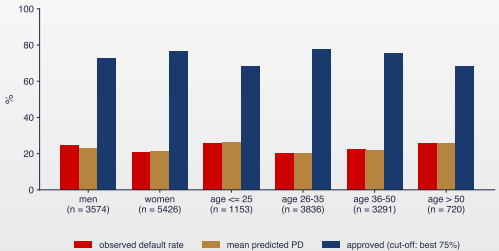
IN THE HOUSE OF REPRESENTATIVES

A BILL

H.R. 10162 (1973), proiectul de lege care a introdus Equal Credit Opportunity Act

Foto: H.R. 10162 (1973); domeniu public; Wikimedia Commons

Aprobarea și nerambursarea pe grupuri: Taiwan



- Eșantionul de test; sexul și vîrsta nu sînt variabile ale modelului; pragul: aprobăm cei 75% cu PD cea mai mică
- Femei: rata de nerambursare 20,5%, aprobate 76,6%; bărbați: 24,7%, 72,5%; bun-platnici aprobați: femei 85,1%, bărbați 83,2%
- Interpretare: modelul este calibrat în fiecare grup (PD medie apropiată de rata observată), deci grupurile cu mai multe nerambursări (bărbații, cei mai tineri și cei mai în vîrstă) sînt aprobate mai rar

SFM_ch12_reject_fairness

Studiu de caz: învățarea automată și inegalitatea accesului la credit

- Fuster et al. (2022): credite ipotecare din SUA; nerambursarea prezisă cu un logit tradițional și cu modele de învățare automată
  - întrebarea: cine cîștigă și cine pierde cînd creditorul trece la modelul mai precis?
- Rezultatele
  - modelele flexibile prezic mai bine nerambursarea
  - debitorii de culoare și cei hispanici au mai puține șanse să cîștige din noua tehnologie
  - într-un model de echilibru, disparitatea dobînzilor între grupuri și în interiorul lor crește, mai ales din cauza flexibilității mai mari
- Lecția: acuratețea și echitatea trebuie măsurate împreună; un AUC mai bun nu este un argument suficient

Scorul Z al lui Altman (1968)

- Altman (1968): 66 de companii industriale americane, 33 care au intrat în faliment în 1946–1965 și 33 de companii pereche care au supraviețuit
  - analiza discriminantă a lui Fisher pe cinci indicatori contabili
- $Z = 1{,}2X_1 + 1{,}4X_2 + 3{,}3X_3 + 0{,}6X_4 + 1{,}0X_5$
  - X_1 capitalul de lucru / activele totale (TA); X_2 profitul reinvestit / TA; X_3 EBIT (profitul înainte de dobînzi și impozite) / TA
  - X_4 valoarea de piață a capitalurilor proprii / valoarea contabilă a datoriilor; X_5 cifra de afaceri / TA
  - zonele: $Z < 1{,}81$ dificultate, 1,81–2,99 zona gri, $Z > 2{,}99$ sigur
- Aproximativ 95% din eșantion clasificat corect cu un an înainte de faliment; acuratețea scade repede pentru orizonturi mai lungi



NYU Stern School of Business, instituția lui Altman

Foto: Ajay Suresh (2019); CC BY 2.0; Wikimedia Commons

Exemplu rezolvat: un scor Z

- O companie ipotetică: $X_1 = 0{,}15$, $X_2 = 0{,}20$, $X_3 = 0{,}08$, $X_4 = 0{,}90$, $X_5 = 1{,}10$
  - $Z = 1{,}2 \times 0{,}15 + 1{,}4 \times 0{,}20 + 3{,}3 \times 0{,}08 + 0{,}6 \times 0{,}90 + 1{,}0 \times 1{,}10$
  - $Z = 0{,}180 + 0{,}280 + 0{,}264 + 0{,}540 + 1{,}100 = 2{,}364$: zona gri
- Ce indicator ar muta compania în zona sigură?
  - un EBIT/TA de 0,27 în loc de 0,08 adaugă $3{,}3 \times 0{,}19 \approx 0{,}63$: $Z \approx 2{,}99$
- Limite
  - ponderile provin din 66 de companii industriale americane din 1946–1965; alte sectoare și țări cer reestimare
  - Z este un scor discriminant, nu o PD; modelele ulterioare dau PD: logit Ohlson (1980), modele de hazard Shumway (2001)

Distanța pînă la nerambursare a lui Merton

- Merton (1974): valoarea V_t a activelor companiei urmează o GBM (mișcarea browniană geometrică, Capitolul 4) cu rata de creștere (drift) μ și volatilitatea σ_V
  - datorie cu valoarea nominală D scadentă la T; nerambursare dacă $V_T < D$; capitalurile proprii sînt o opțiune call pe active
- $\ln V_T \sim N\big(\ln V_0 + (\mu - \sigma_V^2/2)T,\ \sigma_V^2 T\big)$, deci
  - $\mathrm{DD} = \dfrac{\ln(V_0/D) + (\mu - \sigma_V^2/2)T}{\sigma_V\sqrt{T}}, \quad \mathrm{PD} = \Phi(-\mathrm{DD})$
  - DD: numărul de abateri standard dintre valoarea așteptată a activelor și datorie
- V și σ_V nu sînt observabile: se obțin din prețul și volatilitatea acțiunilor (evaluarea opțiunilor)



Robert C. Merton, 2010

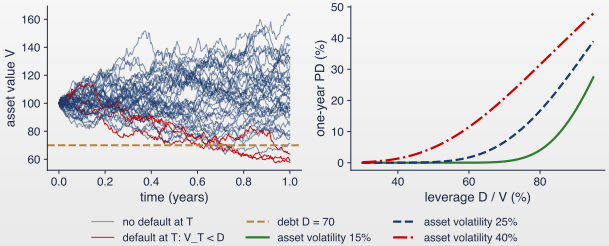
Foto: Massachusetts Institute of Technology (2010); CC BY-SA 4.0; Wikimedia Commons

Exemplu rezolvat: distanța pînă la nerambursare

- $V_0 = 100$, $D = 70$, $\sigma_V = 25\%$, $\mu = 5\%$, $T = 1$ an
  - $\ln(100/70) = 0{,}3567$; $\mu - \sigma_V^2/2 = 0{,}05 - 0{,}03125 = 0{,}01875$
  - $\mathrm{DD} = 0{,}3754/0{,}25 = 1{,}502$; $\mathrm{PD} = \Phi(-1{,}502) = 6{,}66\%$
- Din piață: capitaluri proprii 40, volatilitatea acțiunilor 50%, datorie 70, rata fără risc 3%
  - rezolvînd $E = V\Phi(d_1) - De^{-rT}\Phi(d_2)$ și $\sigma_E E = \Phi(d_1)\sigma_V V$: $V = 107{,}9$, $\sigma_V = 18{,}6\%$, $\mathrm{DD} = 2{,}39$, $\mathrm{PD} = 0{,}84\%$
  - E, σ_E: valoarea și volatilitatea capitalurilor proprii; r: rata fără risc; d_1, d_2: termenii Black–Scholes din Capitolul 4, cu V în locul prețului acțiunii
- Interpretare: o PD din piață reacționează zilnic la prețuri; un scor contabil se schimbă o dată pe an

SFM_ch12_altman_merton

Merton: traiectoriile activelor și PD



- Stînga: 40 de traiectorii simulate ale activelor ($V_0 = 100$, $\mu = 5\%$, $\sigma_V = 25\%$); roșu: traiectoriile care se termină sub datoria de 70
- Dreapta: PD pe un an în funcție de gradul de îndatorare D/V pentru trei volatilități ale activelor
- Interpretare: PD este aproape zero la îndatorare mică și crește abrupt între 70% și 90%; cu cît volatilitatea este mai mare, cu atît creșterea începe mai devreme

SFM_ch12_altman_merton

Studiu de caz: acuratețea predictivă a modelului Merton

- Bharath & Shumway (2008): companii americane; DD Merton față de un DD „naiv", cu aceeași formă funcțională, dar fără rezolvarea modelului
  - metoda: modele de hazard ale nerambursării și prognoze în afara eșantionului; comparație cu spread-urile CDS (credit default swap) și randamentele obligațiunilor
- Rezultatele
  - DD naiv funcționează puțin mai bine decît DD Merton în afara eșantionului
  - alte variabile (de exemplu randamentele trecute) aduc informație: DD nu este o statistică suficientă pentru PD
  - forma sa funcțională (îndatorarea raportată la volatilitate) este ceea ce îl face util
- Lecția: teoria dă o variabilă bună, statistica decide cum o folosim; la fel pentru scorul Z

O întrebare de cercetare deschisă

- **Îmbunătățesc scorurile flexibile de învățare automată prognoza nerambursării suficient pentru a justifica pierderea de transparență și de echitate?**
  - azi: logit-ul WoE și LDA ating un AUC de test de aproximativ 0,764; comparațiile găsesc cîștiguri semnificative pentru ansambluri Lessmann et al. (2015)
  - cîștigurile pot fi inegale între grupuri Fuster et al. (2022); mașini cu vectori suport pentru nerambursare: Chen, Härdle & Moro (2011)
- Întrebarea rămîne deschisă: cîștigurile depind de setul de date, de măsură și de populație; rezultatele solicitanților respinși nu se văd niciodată
- Instrumentele AI pot accelera un astfel de studiu, dar rezultatele lui trebuie verificate Wang et al. (2023)

Utilizări posibile ale instrumentelor AI

- **Literatura**: o listă a comparațiilor de metode de scoring și a studiilor despre echitatea creditării, cu seturile lor de date
- **Cod**: o primă versiune a unei comparații prin validare încrucișată a logit-ului WoE cu gradient boosting pe datele din Taiwan
- **Robustețe**: alte praguri, alte măsuri (Brier, cost), AUC și calibrare pe grupuri
- Exemplu de prompt
  - `Write Python code that compares a WoE logistic regression with gradient boosting on the Taiwan credit card default data by repeated stratified 5-fold cross-validation, recomputing the WoE bins inside each fold, and reports AUC, Brier score and the approval rates of men and women at a 75 percent approval cut-off.`

Verificări necesare

- Codificarea rezultatului: nerambursare $= 1$; o codificare inversă schimbă fiecare semn și fiecare curbă
- Convenția de semn pentru WoE: aici ln(bun/rău); unele surse folosesc ln(rău/bun)
- Fără scurgere de informație: intervalele, WoE și selecția variabilelor în fiecare parte de estimare
- AUC nu este acuratețea, iar Gini $= 2\,\text{AUC} - 1$, nu AUC $-0{,}5$
- PD-urile estimate pe un set de date cu rău-platnici suprareprezentați cer corecția a priori înaintea oricărei cifre de EL sau capital
- Referințele: fiecare lucrare și set de date citat trebuie să existe; verificați DOI-ul

Idee de proiect

- **Întrebarea**: pe datele cardurilor din Taiwan, cît adaugă gradient boosting la AUC față de logit-ul WoE și ce se întîmplă cu ratele de aprobare pe sexe și vîrste?
  - date: Default of Credit Card Clients (2009) (CC BY 4.0), salvate în data/credit
- Pași
  - intervale WoE și IV pentru fiecare variabilă; un logit WoE; gradient boosting cu setările implicite
  - validare încrucișată stratificată repetată: AUC cu intervale DeLong, scorul Brier, calibrarea
  - echitatea la un prag: paritatea demografică și egalitatea de șanse pentru sexe și grupe de vîrstă
- Livrabil: un tabel, un grafic și un paragraf despre ce pot și ce nu pot arăta datele
- Declarați orice utilizare a instrumentelor AI și enumerați erorile acestora pe care le-ați corectat

Idei de reținut

- EL = PD × LGD × EAD intră în preț; capitalul IRB acoperă pierderea neașteptată
- WoE transformă fiecare variabilă într-o dovadă pe scala logaritmului șansei; IV ordonează variabilele
- Logit: e^{β} este un raport al șanselor; LDA (Fisher) dă același logaritm liniar al șansei sub ipoteza de normalitate și ordonează la fel de bine aici
- Un scorecard: Scorul = Offset + Factor ln(șansa bun:rău), PDO = 20 de puncte pentru o dublare
- Validăm discriminarea (AUC, Gini, KS) și calibrarea (Brier, diagrama de calibrare) în afara eșantionului; alegem pragul din costuri
- Solicitanții respinși și grupurile protejate cer verificări explicite; Altman și Merton adaugă informație contabilă și de piață

Formule de reținut

Mărimea	**Formula**
EL	$\mathrm{PD} \times \mathrm{LGD} \times \mathrm{EAD}$
WoE, IV	$\ln \frac{g_i/G}{b_i/B}$, $\sum_i (g_i/G - b_i/B)\,\mathrm{WoE}_i$
Logit	$\ln \frac{p}{1-p} = \beta_0 + x^\top \beta$, $p = 1/(1+e^{-\eta})$
LDA	$w = S_W^{-1}(m_1 - m_0)$
Corecția a priori	$+\ln \frac{\pi/(1-\pi)}{\pi_s/(1-\pi_s)}$ la termenul liber
Scorecard	$\mathrm{Offset} + \frac{\mathrm{PDO}}{\ln 2} \ln \frac{1-p}{p}$
AUC, Gini	$P(s_{\text{rău}} > s_{\text{bun}})$, $2\,\mathrm{AUC} - 1$
KS, Brier	$\max_s \lvert F_{\text{rău}}(s) - F_{\text{bun}}(s)\rvert$, $\frac{1}{n}\sum (p_i - y_i)^2$
Pragul Bayes	$p > C_{FP}/(C_{FP} + C_{FN})$
Merton	$\mathrm{DD} = \frac{\ln(V/D) + (\mu - \sigma^2/2)T}{\sigma\sqrt{T}}$, $\mathrm{PD} = \Phi(-\mathrm{DD})$

Autoevaluare

- **Întrebare**: un model are o acuratețe de 70% pe eșantionul de test South German Credit. Este bun?
  - **Răspuns**: nu neapărat: regula „toți sînt bun-platnici" atinge deja 70%; ne uităm la AUC, TPR, FPR și cost
- **Întrebare**: coeficientul unei variabile indicator într-un logit este 0,69. Ce înseamnă?
  - **Răspuns**: șansa de nerambursare se înmulțește cu $e^{0,69} \approx 2$ pentru grup, față de grupul de referință
- **Întrebare**: schimbă corecția pentru suprareprezentare valoarea AUC?
  - **Răspuns**: nu: deplasează toți logaritmii șansei cu aceeași constantă, deci ordonarea nu se schimbă
- Urmează: Capitolul 13, învățarea automată

Bibliografie (1/2)

- **Altman, E.I. (1968).** Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. ***Journal of Finance***, **23(4), 589–609.**
- **Banasik, J., Crook, J., Thomas, L. (2003).** Sample selection bias in credit scoring models. ***Journal of the Operational Research Society***, **54(8), 822–832.**
- **Basel Committee on Banking Supervision (2005).** An explanatory note on the Basel II IRB risk weight functions. **Bank for International Settlements.**
- **Basel Committee on Banking Supervision (2006).** International convergence of capital measurement and capital standards: a revised framework (comprehensive version). **Bank for International Settlements.**
- **Będowska-Sójka, B., Wójcik, P., Pele, D.T. (2026).** Early warning systems for cryptocurrency markets: predicting 'zombie' assets using machine learning. ***North American Journal of Economics and Finance***, **102543.**
- **Bharath, S.T., Shumway, T. (2008).** Forecasting default with the Merton distance to default model. ***Review of Financial Studies***, **21(3), 1339–1369.**
- **Brier, G.W. (1950).** Verification of forecasts expressed in terms of probability. ***Monthly Weather Review***, **78(1), 1–3.**
- **Chen, S., Härdle, W.K., Moro, R.A. (2011).** Modeling default risk with support vector machines. ***Quantitative Finance***, **11(1), 135–154.**
- **Cox, D.R. (1958).** The regression analysis of binary sequences. ***Journal of the Royal Statistical Society, Series B***, **20(2), 215–232.**
- **Default of Credit Card Clients [data set] (2009).** UCI Machine Learning Repository, **CC BY 4.0.**
- **DeLong, E.R., DeLong, D.M., Clarke-Pearson, D.L. (1988).** Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. ***Biometrics***, **44(3), 837–845.**
- **Durand, D. (1941).** ***Risk Elements in Consumer Instalment Financing***, **technical edition. National Bureau of Economic Research.**
- **European Commission (2024).** AI Act: regulatory framework on artificial intelligence. **Shaping Europe's digital future.**
- **Federal Trade Commission.** Equal Credit Opportunity Act, **15 U.S.C. §§ 1691–1691f.**
- **Fisher, R.A. (1936).** The use of multiple measurements in taxonomic problems. ***Annals of Eugenics***, **7(2), 179–188.**
- **Franke, J., Härdle, W.K., Hafner, C.M. (2019).** ***Statistics of Financial Markets: An Introduction***, **5th ed. Springer.**
- **Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., Walther, A. (2022).** Predictably unequal? The effects of machine learning on credit markets. ***Journal of Finance***, **77(1), 5–47.**

Bibliografie (2/2)

- **Gordy, M.B. (2003).** A risk-factor model foundation for ratings-based bank capital rules. ***Journal of Financial Intermediation***, 12(3), 199–232.
- **Grömping, U. (2019).** South German credit data: correcting a widely used data set. **Reports in Mathematics, Physics and Chemistry 4/2019, Beuth University of Applied Sciences Berlin.**
- **Hand, D.J., Henley, W.E. (1997).** Statistical classification methods in consumer credit scoring: a review. ***Journal of the Royal Statistical Society, Series A***, 160(3), 523–541.
- **Hanley, J.A., McNeil, B.J. (1982).** The meaning and use of the area under a receiver operating characteristic (ROC) curve. ***Radiology***, 143(1), 29–36.
- **Härdle, W.K., Simar, L. (2019).** *Applied Multivariate Statistical Analysis*, **5th ed. Springer.**
- **Hardt, M., Price, E., Srebro, N. (2016).** Equality of opportunity in supervised learning. ***Advances in Neural Information Processing Systems*** **29; arXiv:1610.02413.**
- **Hosmer, D.W., Lemeshow, S. (1980).** Goodness of fit tests for the multiple logistic regression model. ***Communications in Statistics – Theory and Methods***, 9(10), 1043–1069.
- **Lessmann, S., Baesens, B., Seow, H.-V., Thomas, L.C. (2015).** Benchmarking state-of-the-art classification algorithms for credit scoring: an update of research. ***European Journal of Operational Research***, 247(1), 124–136.
- **Merton, R.C. (1974).** On the pricing of corporate debt: the risk structure of interest rates. ***Journal of Finance***, 29(2), 449–470.
- **myFICO. The history of the FICO score. Fair Isaac Corporation.**
- **Ohlson, J.A. (1980).** Financial ratios and the probabilistic prediction of bankruptcy. ***Journal of Accounting Research***, 18(1), 109–131.
- **Shumway, T. (2001).** Forecasting bankruptcy more accurately: a simple hazard model. ***Journal of Business***, 74(1), 101–124.
- **Siddiqi, N. (2006).** *Credit Risk Scorecards: Developing and Implementing Intelligent Credit Scoring*. **Wiley.**
- **South German Credit [data set] (2020). UCI Machine Learning Repository, CC BY 4.0.**
- **Thomas, L.C., Crook, J., Edelman, D. (2017).** *Credit Scoring and Its Applications*, **2nd ed. SIAM.**
- **Wang, H., Fu, T., Du, Y., Gao, W., et al. (2023).** Scientific discovery in the age of artificial intelligence. ***Nature***, 620, 47–60.
- **Yeh, I.-C., Lien, C.-H. (2009).** The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients. ***Expert Systems with Applications***, 36(2), 2473–2480.