

KREDIT TO'LOVINI BAJARMASLIK XAVFINI LSTM MODELI YORDAMIDA BAHOLASH

Sattorov Abdumalik Abduraxmon o'g'li

O'zbekiston Respublikasi Markaziy banki

abdumalikfinance@gmail.com

ARTICLE INFO

Kalit so'zlar:

data science
kredit xavfi
LSTM
vaqt qatori
default
precision
recall
F1
qaror chegarasi

Annotatsiya :

Ushbu maqolada mijozning ketma-ket oylik to'lov tarixidan foydalanib kredit to'lovini bajarmaslik xavfini LSTM (Long Short-Term Memory) modeli orqali baholashning sodda va amaliy tartibi taklif etiladi. Asosiy g'oya shundan iboratki, kredit riski faqat bitta oydagi holat bilan emas, balki to'lov intizomining vaqt bo'yicha qanday o'zgargani bilan ham bog'liq bo'lishi mumkin. Shu sababli olti oylik to'lov holati, qarzdorlikning limitga nisbati va amalga oshirilgan to'lov ulushi ketma-ketlik ko'rinishida modelga uzatiladi. Maqolada LSTM blokining unutish, kirish va chiqish darvozalari formulalari, binary cross-entropy yo'qotish funksiyasi, precision, recall, F1 va xarajatga sezgir qaror chegarasi izohlanadi. Natijalarni tushuntirish uchun 1 000 kuzatuvli shartli sinov, uchta grafik va amaliy hisob-kitob keltiriladi. Grafik va sonlar metodik namuna bo'lib, haqiqiy bank portfeli bo'yicha empirik natija sifatida talqin qilinmaydi.

<http://www.gospodarkainnowacje.pl/> © 2026 LWAB.

Kirish

Kredit to'lovini bajarmaslik xavfini oldindan baholash bankka tekshirish resurslarini to'g'ri taqsimlash, muammoli kreditlarga erta e'tibor qaratish va mijoz bilan ishlash ustuvorligini belgilashda yordam beradi. Oddiy jadval modellari mijozning ma'lum bir sanadagi holatini yaxshi ifodalaydi, biroq to'lov intizomidagi ketma-ket o'zgarishlarni alohida tayyorlangan belgilar orqali ifodalash talab qilinadi. LSTM modeli esa vaqt bo'yicha ketma-ket ma'lumotni bevosita qayta ishlashga mo'ljallangan rekurrent neyron tarmoq turidir [1].

Maqolaning maqsadi LSTMni murakkab nazariy konstruktsiya sifatida emas, bankdagi kredit xavfi vazifasiga mos, tekshiriladigan va tushunarli jarayon sifatida bayon qilishdir. Asosiy e'tibor ma'lumotni to'g'ri shakllantirish, kelajak ma'lumotining sizib chiqishiga yo'l qo'ymaslik, model

natijasini bank uchun foydali ko'rsatkichlar bilan baholash va qaror chegarasini xato qiymatlariga mos tanlashga qaratiladi.

1. Muammo qo'yilishi va ma'lumotlar tuzilishi

Tadqiqot vazifasida maqsadli o'zgaruvchi $y \in \{0,1\}$ ko'rinishida belgilanadi: $y=1$ — keyingi nazorat davrida kredit to'lovini bajarmaslik yoki bank belgilagan default mezoniga tushish; $y=0$ — default kuzatilmaligi. Prognoz sanasi t_0 deb olinadi va modelga faqat t_0 dan oldin mavjud bo'lgan ma'lumot kiritiladi. Bu qoida real amaliyotda eng muhim metodik cheklovlardan biri bo'lib, aks holda model kelajakdagi natijani bilib olgan holda sun'iy ravishda yuqori sifat ko'rsatishi mumkin.

Ochiq tajriba uchun UCI Machine Learning Repository'dagi Default of Credit Card Clients to'plami mos boshlang'ich manba bo'la oladi. Unda 30 000 mijoz bo'yicha kredit limiti, oldingi oylar to'lov holati, hisoblangan qarzdorlik va amalda to'langan summalar mavjud [2]. Mazkur maqoladagi yondashuvda bu ustunlar oddiy jadval ko'rinishida emas, olti oylik ketma-ketlik sifatida qayta tashkil qilinadi.

1-jadval. Ketma-ketlik uchun tavsiya etiladigan asosiy belgilar

Belgi	Hisoblash / mazmuni	LSTMdagi roli
To'lov holati PAY _t	Har oy kechikish holati kodi	Intizomning vaqt bo'yicha o'zgarishi
Qarz/limit nisbati	BILL _t / LIMIT	Limitdan foydalanish bosimini ko'rsatadi
To'lov ulushi	PAY_AMT _t / max(BILL _t ,1)	Qarzning qanchasi qoplanganini ifodalaydi
Qarz o'sishi	(BILL _t -BILL _{t-1}) / LIMIT	Qarzdorlik dinamikasini ifodalaydi

Har bir mijoz uchun $X = [x_1, x_2, \dots, x_6]$ ketma-ketligi hosil qilinadi. Bu yerda x_t — t -oydagi uch yoki to'rtta moliyaviy belgidan tashkil topgan vektor. Masalan, $x_t = [\text{PAY}_t, \text{BILL}_t/\text{LIMIT}, \text{PAY_AMT}_t/\text{BILL}_t, \Delta\text{BILL}_t/\text{LIMIT}]$. Shunday qilib, LSTM bir vaqtning o'zida nafaqat ko'rsatkich darajasini, balki ularning qaysi oyda va qanday ketma-ketlikda yuz berganini ham ko'radi.

$$X_i = [x_{(i,1)}, x_{(i,2)}, \dots, x_{(i,6)}], \quad x_{(i,t)} \in \mathbb{R}^k \quad (1)$$

Kategoriya kodlari va nolga bo'lish holatlari oldindan qat'iy qoidalar bilan qayta ishlanadi. Nisbatlarda $\max(\text{BILL}_t, 1)$ kabi himoya ishlatilishi mumkin. Juda katta qiymatlar winsorization yoki log-transformatsiya orqali yumshatiladi, biroq transformatsiya chegaralari faqat o'qitish qismida aniqlanadi. Mijoz identifikatori modelga belgi sifatida berilmaydi.[3]

Agar bir mijoz bo'yicha bir nechta prognoz sanalari mavjud bo'lsa, tasodifiy ajratish o'rniga vaqt bo'yicha ajratish ma'qul: oldingi davrlar o'qitish, keyingi davrlar validatsiya va test uchun olinadi. Bu bank amaliyotidagi haqiqiy holatga yaqinroq baholash beradi.

2. Ma'lumotlarni tayyorlash va ketma-ketlik yaratish

LSTM uchun ma'lumot tayyorlashning eng qulay shakli uch o'lchamli massivdir: [mijozlar soni, oylar soni, belgilar soni]. Masalan, 20 000 ta o'qitish kuzatuv, 6 oy va 4 ta vaqtga bog'liq belgi ishlatilsa, kirish tenzori o'lchami $20\,000 \times 6 \times 4$ bo'ladi. Kredit limiti, yosh yoki shartnoma turi kabi vaqt bo'yicha o'zgarmaydigan belgilar alohida vektor sifatida saqlanib, LSTM chiqishi bilan keyinroq birlashtirilishi mumkin. Soddalashtirilgan variantda esa faqat ketma-ket belgilar bilan ishlash yetarli.[4]

2-jadval. LSTM uchun ma'lumot tayyorlash tartibi

Bosqich	Amal	Nazorat nuqtasi
1	Prognoz sanasini belgilash	Kelajak ustunlari kiritilmasligi
2	Oldingi 6 oyni tanlash	Har mijozda bir xil vaqt oynasi
3	Nisbat belgilarini hisoblash	Nol va chet qiymatlar nazorati
4	Train/validation/test ajratish	Vaqt yoki mijoz bo'yicha ajratish
5	Masshtablash	Parametrlar faqat train'da aniqlanadi
6	Tensor shakliga keltirish	[N, 6, k] o'lcham tekshiruvi

Masshtablash LSTMning o'qitilishini barqarorlashtiradi. Masalan, z-score usulida har bir uzluksiz belgi o'qitish qismidagi o'rtacha qiymat μ va standart og'ish σ yordamida standartlashtiriladi:

$$z = (x - \mu_{\text{train}}) / \sigma_{\text{train}} \quad (2)$$

Muhim talab shundaki, validatsiya va test to'plamlarining o'rtacha qiymati yoki standart og'ishi masshtablash parametrlarini hisoblashda ishlatilmaydi. Aks holda test ma'lumoti bilvosita o'qitish jarayoniga kirib qoladi. Missing qiymatlar uchun ham xuddi shu tamoyil amal qiladi: to'ldirish qoidasi train qismida aniqlanadi va qolgan qismlarga o'zgarmas holda qo'llanadi.

Sinflar nomutanosib bo'lishi mumkin. Masalan, default ulushi 20% bo'lsa, barcha mijozni 'default yo'q' deb belgilagan sodda model 80% accuracy beradi. Shu sababli LSTMni baholashda faqat accuracy yetarli emas. Model o'qitishda class_weight yoki focal loss kabi usullar qo'llanishi mumkin, biroq mazkur maqolada tushunarli va sodda bo'lishi uchun standart binary cross-entropy hamda keyingi bosqichda qaror chegarasini sozlash yondashuvi tanlanadi.[5]

Ketma-ketlik uzunligini olti oy deb olish amaliy soddalashtirishdir. U bank mahsulotiga qarab 3, 6, 12 yoki undan ko'p oy bo'lishi mumkin. Uzunlik validatsiya natijasi va biznes mazmuniga ko'ra tanlanadi. Ko'proq oy har doim ham yaxshi natija bermaydi: juda uzun tarix eski va kam ahamiyatli signallarni ham modelga olib kirishi mumkin.

3. LSTM modelining matematik asosi

Oddiy rekurrent neyron tarmoq ketma-ketlikning oldingi holatini keyingi bosqichga uzatadi, lekin uzoq bog'liqliklarda gradientning yo'qolishi yoki ortib ketishi muammosiga duch kelishi mumkin. LSTM bu muammoni maxsus xotira holati va uchta asosiy darvoza orqali yumshatadi [6]. Har bir t vaqt bosqichida kirish x_t , avvalgi yashirin holat h_{t-1} va avvalgi xotira c_{t-1} ishlatiladi.

Unutish darvozasi oldingi xotiraning qaysi qismi saqlanishini belgilaydi:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (3)$$

Kirish darvozasi yangi ma'lumotning qaysi qismi xotiraga yozilishini boshqaradi. Shu bilan birga yangi nomzod xotira vektori hosil qilinadi:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (4)$$

$$\tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (5)$$

Yangilangan xotira oldingi xotira va yangi nomzodning og'irlangan yig'indisi sifatida hisoblanadi:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (6)$$

Chiqish darvozasi esa joriy yashirin holatni hosil qiladi:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (7)$$

$$h_t = o_t \odot \tanh(c_t) \quad (8)$$

Bu formulalarda σ — sigmoid funksiya, $\tanh$ — giperbolik tangens, $\odot$ — elementlar bo'yicha ko'paytirish. Bank ma'lumotida f_t eski oylar haqidagi ayrim axborotni saqlash yoki unutishga, i_t yangi oydagi kechikish va qarzdorlik signalini xotiraga kiritishga, o_t esa keyingi qatlamga uzatiladigan umumlashtirilgan holatni shakllantirishga xizmat qiladi.

Olti oylik ketma-ketlik qayta ishlangach, oxirgi yashirin holat h_6 to'liq bog'langan (Dense) qatlamga beriladi va sigmoid orqali default skori olinadi:

$$\hat{p} = \sigma(w^T h_6 + b) \quad (9)$$

$\hat{p}$ 0 va 1 oralig'idagi skor bo'lib, kalibrlash tekshiruviz uni to'g'ridan-to'g'ri 'aniq default ehtimoli' deb qabul qilish to'g'ri emas. Amaliy qaror uchun $\hat{p}$ belgilangan τ chegarasi bilan solishtiriladi: $\hat{p} \geq \tau$ bo'lsa 1-sinf, aks holda 0-sinf. Shu sababli model sifatini faqat neyron tarmoq arxitekturasini emas, τ chegarasining tanlanishi ham belgilaydi.[7]

4. Model arxitekturasini va o'qitish tartibi

Sodda boshlang'ich arxitektura sifatida 6×4 o'lchamli ketma-ketlik, 32 ta yashirin birlikli bitta LSTM qatlami, 0,2 dropout va bitta sigmoid chiqish neyroni yetarli. Bu konfiguratsiya katta transformerlarga yoki ko'p qatlamli tarmoqlarga qaraganda tushunarli, tez sinovdan o'tadi va kichikroq ma'lumotda ortiqcha murakkablik xavfini kamaytiradi. Keyingi tajribada 16, 32 va 64 birlikli variantlar validatsiyada taqqoslanishi mumkin.

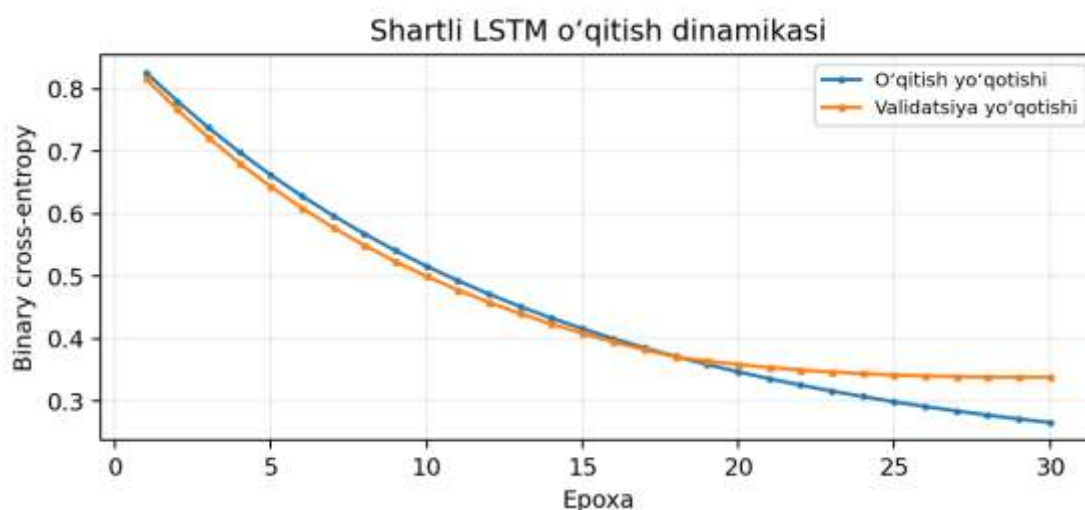
3-jadval. Tavsiya etiladigan sodda LSTM konfiguratsiyasi[8]

Qism	Boshlang'ich qiymat	Izoh
Kirish	6 oy $\times$ 4 belgi	Har mijoz uchun ketma-ketlik
LSTM units	32	16/32/64 validatsiyada solishtiriladi
Dropout	0,20	Haddan tashqari moslashishni kamaytirish
Chiqish	1 sigmoid	Default skori $\hat{p}$
Optimizer	Adam, lr=0,001	Standart boshlang'ich variant
Batch size	64	Xotira va tezlikka mos
Epoch	30 gacha	Early stopping bilan

Ikki sinfli vazifa uchun binary cross-entropy yo'qotish funksiyasi ishlatiladi:

$$L = - (1/N) \sum [y_i \ln(\hat{p}_i) + (1-y_i) \ln(1-\hat{p}_i)] \quad (10)$$

O'qitish jarayonida har bir epoxadan so'ng validatsiya yo'qotishi kuzatiladi. Agar u ketma-ket bir necha epoxa davomida yaxshilanmasa, early stopping qo'llanadi va eng yaxshi validatsiya holatidagi vaznlar qayta tiklanadi. Bu yondashuv modelni train ma'lumotiga ortiqcha moslashishdan himoya qiladi.[9]



1-rasm. Shartli misolda o'qitish va validatsiya yo'qotishining o'zgarishi. Grafik metodik tasvir bo'lib, haqiqiy model natijasi emas.

- 1- rasmdagi shartli trayektoriyada train yo'qotishi doimiy pasayadi, validatsiya yo'qotishi esa taxminan 18–22-epoxalardan keyin sekinlashadi. Amaliyotda shu nuqtadan keyingi keskin farq overfitting signali bo'lishi

mumkin. Model tanlashda yakuniy test to'plamiga qarab epoxa sonini belgilash mumkin emas; test faqat yakuniy, bir martalik baholash uchun saqlanadi.

5. Baholash mezonlari va shartli hisob-kitob

Default 1-sinf deb olinadi. TP — default bo'lgan va model ham default deb belgilagan holat; FN — default bo'lgan, ammo model aniqlamagan holat; FP — default kuzatilmagan mijozga noto'g'ri xavf signali; TN — default bo'lmagan va to'g'ri aniqlangan holat. Bank uchun FN ko'pincha muhim, chunki xavfli mijoz e'tibordan chetda qoladi. FP esa ortiqcha tekshiruv xarajatini keltirib chiqaradi.[10]

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \quad (11)$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \quad (12)$$

$$\text{F1} = 2 \cdot \text{Precision} \cdot \text{Recall} / (\text{Precision} + \text{Recall}) \quad (13)$$

$$\text{Accuracy} = (\text{TP} + \text{TN}) / N \quad (14)$$

4-jadval. 1 000 kuzatuvli shartli LSTM xatolar matritsasi ($\tau=0,50$)

Haqiqiy holat	Prognoz: default (1)	Prognoz: default yo'q (0)
Default (1)	TP = 156	FN = 44
Default yo'q (0)	FP = 96	TN = 704

Ushbu shartli jadvalda jami 200 ta default va 800 ta default bo'lmagan kuzatuv mavjud. Hisob-kitoblar quyidagicha:

$\text{Precision} = 156 / (156 + 96) = 156 / 252 = 0,6190$, ya'ni 61,90%. Bu model default deb belgilagan mijozlarning taxminan 62 foizi haqiqatan ham default sinfiga tegishli ekanini bildiradi.

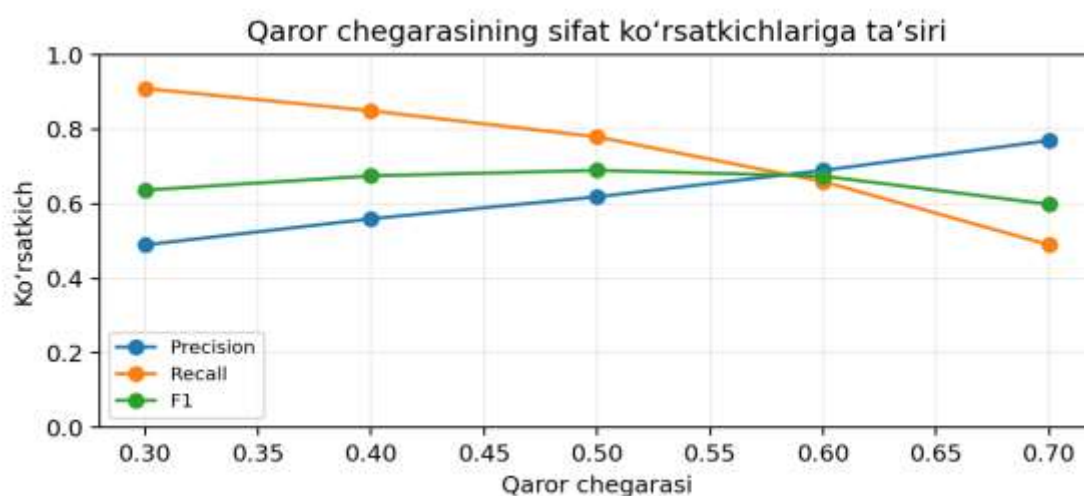
$\text{Recall} = 156 / (156 + 44) = 156 / 200 = 0,78$, ya'ni 78%. Demak, mavjud default holatlarining 78 foizi aniqlangan, 22 foizi esa aniqlanmay qolgan.

$\text{F1} = 2 \times 0,6190 \times 0,78 / (0,6190 + 0,78) \approx 0,6903$, ya'ni 69,03%. $\text{Accuracy} = (156 + 704) / 1000 = 86\%$. Faqat barcha kuzatuvni 'default yo'q' deb belgilaydigan bazaviy qoida 80% accuracy beradi, lekin recall 0% bo'ladi. Shuning uchun 86% accuracy alohida olinganda model foydaliligini to'liq ifodalamaydi.

Natijalarni talqin qilishda bazaviy model bilan bir xil test to'plamidagi taqqoslash shart. Agar LSTM recall yoki xarajat mezonida sodda logistic regression yoki Decision Tree'dan ustun bo'lmasa, murakkabroq modeldan foydalanishning amaliy asosi zaiflashadi. LSTMning afzalligi ayniqsa ketma-ketlikdagi naqshlardan qo'shimcha signal olish bilan namoyon bo'lishi kerak.[11]

6. Qaror chegarasi va sifat ko'rsatkichlari

Sigmoid chiqishi bo'yicha standart $\tau=0,50$ chegarasi har doim bank uchun optimal emas. Chegara pasaytirilsa, ko'proq mijoz xavfli deb belgilanadi: recall oshadi, lekin false positive soni ham ko'payadi. Chegara ko'tarilsa, ogohlantirishlar aniqroq bo'lishi mumkin, ammo ayrim haqiqiy default holatlari o'tkazib yuboriladi. Shu sababli qaror chegarasi validatsiya ma'lumotida bankning maqsadiga mos tanlanadi.[12]



2- rasm. Shartli misolda qaror chegarasi o'zgarganda precision, recall va F1 dinamikasi.

5-jadval. Qaror chegarasi bo'yicha shartli sifat ko'rsatkichlari[13]

τ	Precision	Recall	F1
0,30	49%	91%	63,7%
0,40	56%	85%	67,5%
0,50	61,9%	78%	69,0%
0,60	69%	66%	67,5%
0,70	77%	49%	59,9%

2-rasm va 5-jadvalda $\tau=0,50$ atrofida F1 eng yuqori qiymatga yaqinlashmoqda, lekin bu avtomatik ravishda eng yaxshi bank qarori degani emas. Masalan, bank uchun defaultni o'tkazib yuborish false positive'dan besh baravar qimmat bo'lsa, recallga ko'proq ustuvorlik berish kerak bo'ladi. Shu sababli F1 bilan birga xarajatga sezgir mezon ham hisoblanadi.

Chegara tanlash jarayoni uch bosqichda amalga oshiriladi: birinchidan, LSTM arxitekturasini va o'qitish parametrlari train/validation orqali tanlanadi; ikkinchidan, validatsiyada bir nechta τ qiymati bo'yicha precision, recall, F1 va xarajat solishtiriladi; uchinchidan, tanlangan model va τ kombinatsiyasi o'zgartirilmasdan yakuniy testda bir marta tekshiriladi. Bu tartib test natijasiga moslab modelni yashirin ravishda sozlash xavfini kamaytiradi.

Kreditni avtomatik rad etish uchun faqat LSTM skoridan foydalanish tavsiya etilmaydi. Skor qo'shimcha tekshiruv, collection strategiyasi yoki risk-monitoring navbatini belgilashda yordam beruvchi signal sifatida qo'llanishi mumkin. Yakuniy biznes qoida ichki siyosat, qonunchilik, mijoz huquqlari va model riskini boshqarish talablari bilan uyg'unlashtirilishi zarur.

7. Xatolar qiymati bo'yicha optimal yechim

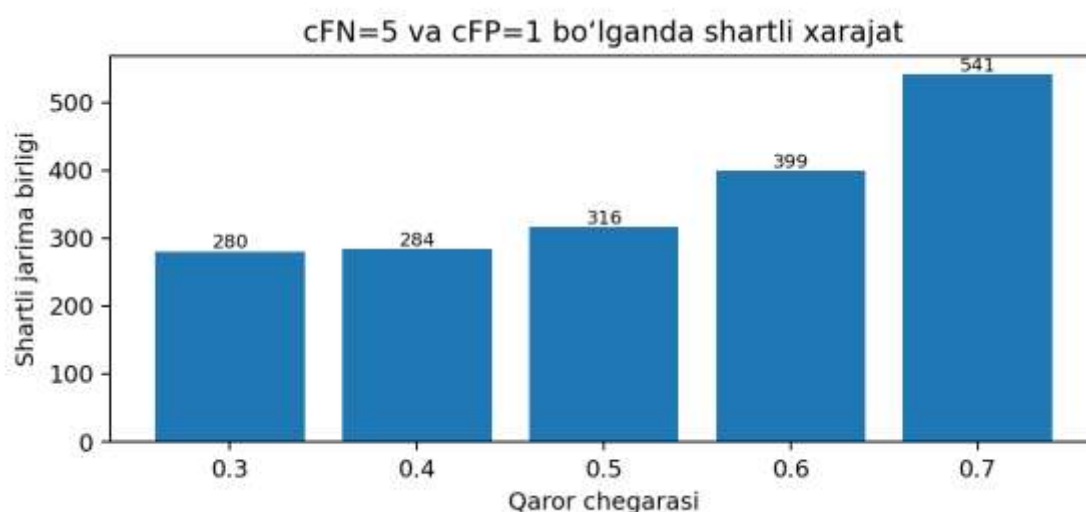
Bank nuqtai nazaridan ikki turdagi xatoning iqtisodiy oqibati bir xil emas. FN holatida potentsial muammoli kredit aniqlanmay qoladi; FP holatida esa yaxshi mijoz ortiqcha tekshiruvga tushadi. Soddalashtirilgan baholash uchun cFN va cFP koeffitsiyentlari kiritiladi:

$$C(\tau) = c_{FN} \cdot FN(\tau) + c_{FP} \cdot FP(\tau) \quad (15)$$

Misolda cFN=5 va cFP=1 shartli birlik deb olinsin. Bu raqamlar haqiqiy pul zarari emas, faqat defaultni o'tkazib yuborishning nisbatan qimmatroq ekanini ifodalaydi. Turli chegaralarda shartli FN va FP sonlari asosida xarajat quyidagicha bo'ladi.[14]

6-jadval. Shartli xarajat mezonni bo'yicha chegaralarni taqqoslash

τ	FN	FP	$C = 5 \cdot FN + FP$
0,30	18	190	280
0,40	30	134	284
0,50	44	96	316
0,60	68	59	399
0,70	102	31	541

*3-rasm. cFN=5 va cFP=1 bo'lganda qaror chegarasi bo'yicha shartli xarajat.*

Bu misolda eng kichik C qiymati $\tau=0,30$ da 280 birlikni tashkil etadi. $\tau=0,40$ juda yaqin natija — 284 birlik beradi. F1 bo'yicha $\tau=0,50$ nisbatan yaxshi ko'ringan bo'lsa ham, bank defaultni o'tkazib yuborishga katta jarima qo'yganda pastroq chegara ma'qul bo'lishi mumkin. Demak, 'eng yaxshi model' tushunchasi faqat bitta statistik ko'rsatkich bilan emas, biznes xarajati bilan birga aniqlanadi.

Amaliy bankda cFN va cFP kredit summasi, undirish ehtimoli, tekshiruv xarajati va mijozga xizmat ko'rsatish oqibatlarini hisobga olgan holda segmentlar bo'yicha farqlanishi mumkin. Biroq maqola doirasida bitta sodda koeffitsiyentli formula metodik g'oyani tushuntirish uchun yetarli. Muhim jihat — xarajat koeffitsiyentlari test natijasiga qarab keyin o'zgartirilmasligi, balki biznes talabi asosida oldindan belgilanishidir.

8. Amaliy joriy etish, nazorat va tushuntirish

LSTM modelini bank tizimiga joriy etish faqat modelni o'qitish bilan tugamaydi. Birinchi navbatda ma'lumot oqimi barqaror bo'lishi kerak: har oy bir xil formatdagi to'lov holati, qarzдорlik, to'lov summasi va limit ma'lumoti yig'iladi. Feature pipeline prognoz sanasidan keyingi ma'lumotni avtomatik ravishda chiqarib tashlashi kerak. Model versiyasi, ishlatilgan ma'lumot oralig'i va qaror chegarasi audit uchun qayd etiladi.

Ikkinchi muhim masala monitoringdir. Model ishlab turgan davrda default ulushi, asosiy belgilar taqsimoti, prediction score taqsimoti, precision va recall vaqt bo'yicha kuzatiladi. Agar iqtisodiy sharoit yoki kredit portfeli tarkibi sezilarli o'zgarsa, modelning eski naqshlari yangi mijozlar uchun mos kelmasligi mumkin. Bunday holatda qayta o'qitish yoki modelni yangilash masalasi ko'rib chiqiladi.

7-jadval. LSTM modelini ishlab turish davridagi asosiy monitoring

Nazorat yo'nalishi	Kuzatiladigan ko'rsatkich	Amaliy harakat
Ma'lumot sifati	Missing, diapazon, kodlar	ETL xatosini aniqlash
Populyatsiya siljishi	Belgi va score taqsimoti	PSI/taqqoslash, qayta tekshiruv
Model sifati	Precision, recall, F1, AUC	Chegarani yoki modelni qayta ko'rish
Biznes natijasi	FN/FP xarajati	Risk siyosatini moslashtirish
Adolatlilik	Segmentlar bo'yicha xatolar	Kamsitish xavfini tekshirish

LSTMning Decision Tree'ga nisbatan asosiy kamchiligi — tushuntirishning murakkabroq ekanidir. Daraxtda aniq qoidalar ko'rinadi, LSTMda esa yuzlab parametrlar birgalikda ishlaydi. Shu sababli amaliy joriy etishda SHAP, permutation importance yoki vaqt bosqichlari bo'yicha sensitivity tahlili kabi usullar qo'shimcha ishlatilishi mumkin. Biroq bunday tushuntirishlar sababiy xulosa emas; ular model prognoziga qaysi belgilar ko'proq ta'sir qilganini tavsiflaydi.

Soddalashtirilgan amaliy yechim sifatida LSTM skori va an'anaviy jadval modeli parallel ishlatilishi mumkin. Agar ikkala model ham yuqori xavf ko'rsatsa, tekshiruv ustuvorligi oshiriladi; agar natijalar keskin farq qilsa, qo'shimcha tekshiruvga yuboriladi. Bu 'champion-challenger' yondashuvi yangi modelning foydasini nazoratli tarzda baholashga yordam beradi.

Mazkur metodikaning cheklovi shundaki, maqoladagi grafik va hisoblar shartli. Ular haqiqiy bank ma'lumotlarida modelning aniq samaradorligini isbotlamaydi. Empirik xulosa chiqarish uchun real ma'lumotda vaqt bo'yicha test, segmentlar tahlili, kalibrlash va kamida bir nechta bazaviy model bilan taqqoslash talab etiladi.[15]

Xulosa

Taklif etilgan yondashuv kredit xavfini olti oylik to'lov dinamikasi asosida baholashga qaratilgan. LSTMning asosiy afzalligi — kechikish, qarzдорlik va to'lov ulushining ketma-ket o'zgarishlarini alohida qo'lda trend belgilarini ko'p yaratmasdan umumlashtira olishidir. Shu bilan birga model sifati ma'lumot tayyorlash qoidalari, vaqt bo'yicha to'g'ri ajratish, validatsiya va biznesga mos qaror chegarasiga kuchli bog'liq.

Shartli hisobda $\tau=0,50$ uchun precision 61,90%, recall 78%, F1 69,03% va accuracy 86% olindi. Bu sonlar metodik tushuntirish uchun keltirildi. Xarajat mezonini cFN=5 va cFP=1 deb olinganda $\tau=0,30$ shartli xarajatni eng kichik qiymatga tushirdi. Mazkur misol statistik jihatdan eng yuqori F1 va biznes uchun eng maqbul qaror har doim bir xil bo'lmashligini ko'rsatadi.

Amaliy yechim quyidagicha umumlashtiriladi: prognoz sanasidan oldingi olti oylik ma'lumotdan ketma-ketlik tuziladi; ma'lumot vaqt bo'yicha train/validation/test qismlariga ajratiladi; 32 birlikli sodda LSTM early stopping bilan o'qitiladi; validatsiyada model va τ chegarasi tanlanadi; test qismida precision, recall,

F1, xatolar matritsasi va xarajat mezoni bir martalik baholanadi; yakuniy skor kreditni avtomatik rad etuvchi yagona omil emas, balki risk-monitoring va qo'shimcha tekshiruv uchun signal sifatida qo'llanadi. Xulosa qilib aytganda, LSTM kredit to'lovi tarixida vaqt bo'yicha bog'liqlik kuchli bo'lgan vazifalarda foydali model bo'lishi mumkin. Biroq u avtomatik ravishda Decision Tree yoki logistic regression'dan yaxshiroq degani emas. Modelning haqiqiy qiymati bir xil test sharoitida sodda bazaviy modellar bilan taqqoslanganda, defaultlarni yaxshiroq aniqlash, iqtisodiy xarajatni kamaytirish va natijani barqaror saqlash orqali isbotlanishi kerak. Keyingi ishda ushbu metodikani ochiq UCI ma'lumotida to'liq o'qitish tajribasi, ROC-AUC/PR-AUC, kalibrlash va SHAP tahlili bilan kengaytirish maqsadga muvofiq.

Foydalanilgan adabiyotlar:

- [1] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.
- [2] I.-C. Yeh and C.-H. Lien, "The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients," *Expert Systems with Applications*, vol. 36, no. 2, pp. 2473–2480, 2009, doi: 10.1016/j.eswa.2007.12.020.
- [3] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016, 775 p.
- [4] F. Chollet, *Deep Learning with Python*, 2nd ed. Shelter Island, NY, USA: Manning Publications, 2021, 504 p.
- [5] TensorFlow, "LSTM layer," *TensorFlow Keras Documentation*. [Online]. Available: TensorFlow official documentation.
- [6] scikit-learn, "Metrics and scoring: Quantifying the quality of predictions," *scikit-learn Documentation*. [Online]. Available: scikit-learn official documentation.
- [7] scikit-learn, "Tuning the decision threshold for class prediction," *scikit-learn Documentation*. [Online]. Available: scikit-learn official documentation.
- [8] T. Saito and M. Rehmsmeier, "The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets," *PLoS ONE*, vol. 10, no. 3, Art. no. e0118432, 2015, doi: 10.1371/journal.pone.0118432.
- [9] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems 30 (NeurIPS 2017)*, pp. 4765–4774, 2017.
- [10] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.
- [11] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006, doi: 10.1016/j.patrec.2005.10.010.
- [12] J. Davis and M. Goadrich, "The relationship between Precision-Recall and ROC curves," in *Proceedings of the 23rd International Conference on Machine Learning (ICML)*, Pittsburgh, PA, USA, 2006, pp. 233–240, doi: 10.1145/1143844.1143874.
- [13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.
- [14] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, San Diego, CA, USA, 2015, pp. 1–15.
- [15] M. A. Hearst, S. T. Dumais, E. Osman, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intelligent Systems*, vol. 13, no. 4, pp. 18–28, 1998, doi: 10.1109/5254.708428.

