SUST SHAKLLANGAN JARAYONLARNI BASHORATLASHDA INTELLEKTUAL ALGORITMLAR MAJMUASI VA ROBAST KOMBINATSIYALASH METODOLOGIYASI

Pulatov G‘iyos Gofurjonovich[1], Pulatova Gulxayo Azamjon qizi[2]

[1]“TIQXMMI” Milliy tadqiqot universiteti
[2]Farg‘ona davlat texnika universiteti
E-mail: gulxayopulatova1995@gmail.comgiyospulatov1987@gmail.com

K E Y W O R D S

sust shakllangan jarayonlar, ansambl bashoratlash, MLP, ANFIS, LSTM, GRU, robast optimallashtirish

A B S T R A C T

Ushbu maqolada sust shakllangan jarayonlarni, xususan yurak-qon tomir kasalliklari bilan bog‘liq bemorlar murojaati chastotasini, ko‘p o‘lchamli vaqt qatori sifatida bashoratlash masalasi ko‘rib chiqilgan. To‘rtta kontseptual jihatdan farqli bazaviy intellektual algoritm - MLP, ANFIS, LSTM va GRU - qurilgan va real meteorologik-geliogeofizik ma’lumotlar to‘plamida sinovdan o‘tkazilgan. Bazaviy modellar bashoratlarini birlashtirish uchun robast optimallashtirish nazariyasiga asoslangan adaptiv kombinatsiya algoritmi taklif etilgan, u modellar bashoratlari orasidagi yuqori korrelyatsiya sharoitida ham og‘irliklarni turg‘un va shovqinga bardoshli tarzda aniqlaydi. Eksperiment natijalari shuni ko‘rsatdiki, taklif etilgan robast kombinatsiya eng yaxshi bazaviy modeldan past RMSE ko‘rsatkichiga erishadi. Olingan natijalar ansambl yondashuvining sust shakllangan jarayonlarni bashoratlashdagi amaliy samaradorligini tasdiqlaydi.

1. Kirish

Ko‘p o‘lchamli vaqt qatorlari ko‘rinishida namoyon bo‘ladigan sust shakllangan jarayonlarni - nostatsionar, nochiziqli, ko‘p omilli va noaniqlikka ega dinamik jarayonlarni - yuqori aniqlikda bashoratlash zamonaviy sun’iy intellekt tadqiqotlarining dolzarb yo‘nalishlaridan biri hisoblanadi [1, 2]. Bunday jarayonlarga misol sifatida meteorologik va geliogeofizik omillarga sezgir bo‘lgan yurak-qon tomir kasalliklari bilan bog‘liq bemorlar murojaati chastotasini keltirish mumkin: Jahon sog‘liqni saqlash tashkiloti ma’lumotlariga ko‘ra, yurak-qon tomir kasalliklari har yili dunyoda taxminan 19,8 million kishi hayotiga zomin bo‘lmoqda [3].

Yakka intellektual bashoratlash algoritmlari - sun’iy neyron tarmoqlar, neyro-noravshan modellar - o‘z afzalliklariga ega bo‘lsa-da, ularning bashorat xatosi vaqt bo‘yicha barqaror emasligi ko‘plab tadqiqotlarda qayd etilgan [4, 5]. Bir vaqt oralig‘ida bir model aniqroq natija bersa, boshqa oraliqda boshqa model ustunlik qiladi. Bu holat kontseptual jihatdan farqli algoritmlarni ansambl asosida birlashtirishni zarur qiladi [6].

Ansambl bashoratlashning eng oddiy usuli - bazaviy modellar natijalarining oddiy yoki vaznli o‘rtachasini olish [7]. Biroq modellar soni ortgan va ularning bashoratlari o‘zaro yuqori korrelyatsiyaga ega bo‘lgan sharoitda klassik eng kichik kvadratlar usuli bilan optimal og‘irliklarni aniqlash masalasi noturg‘un tus oladi [8]. Ushbu muammoni hal etishning samarali yo‘li A.N. Tixonov tomonidan asoslangan regulyarizatsiya nazariyasidir [8, 9]. So‘nggi yillarda ushbu g‘oyaning ansambl bashoratlashga tatbiq etilishi bo‘yicha diqqatga sazovor natijalar e’lon qilindi: xususan, Bertsimas va Boussioux [10] adaptiv robast optimallashtirish yondashuvini taklif etib, havo sifati, energiya iste’moli va tropik siklonlar intensivligi ma’lumotlarida ansambldagi eng yaxshi yakka modeldan RMSE bo‘yicha 16–26% gacha ustunlikka erishgan.

Adabiyotlar tahlili shuni ko‘rsatadiki, ansambl bashoratlash sohasida ikki yo‘nalish alohida rivojlanmoqda: bir tomondan, chuqur

(cc) BY

o‘qitish arxitekturalarining o‘zini takomillashtirish [4, 5], ikkinchi tomondan - allaqachon mavjud bo‘lgan kontseptual jihatdan farqli modellarni matematik jihatdan asosli tarzda birlashtirish usullarini ishlab chiqish [6, 7, 10]. Klassik bagging va boosting texnikalari asosan gomogen bazaviy modellar ustida samarali ishlaydi, geterogen - ya’ni statik, noravshan-mantiqiy va rekurrent tabiatga ega modellarni birlashtirish uchun esa alohida matematik apparat talab etiladi, chunki bunday modellarning bashorat xatolari orasidagi statistik bog‘liqlik tuzilishi murakkab bo‘ladi. Aynan shu muammoni yechishda Tixonov tipidagi regulyarizatsiya nazariy jihatdan mos vosita hisoblanadi, chunki u dastlab noturg‘un teskari masalalarni yechish uchun ishlab chiqilgan bo‘lib, o‘zaro korrelyatsiyalangan kirish o‘zgaruvchilari sharoitida ham yagona va barqaror yechim topishga kafolat beradi [8, 9].

Ushbu maqolaning maqsadi – sust shakillangan jarayonlarni bashoratlash uchun to‘rtta kontseptual jihatdan farqli bazaviy intellektual algoritmdan iborat majmua yaratish hamda ularning bashoratlarini Tixonov robast optimallashtirish usuliga asoslangan holda adaptiv tarzda birlashtiruvchi kombinatsiya algoritmini ishlab chiqish va uning samaradorligini yurak-qon tomir kasalliklari bilan bog‘liq bemorlar murojaati misolida empirik baholashdir. Tadqiqotning vazifalari quyidagilardan iborat: (1) to‘rtta bazaviy IA ni sirpanib boruvchi oyna texnikasi asosida qurish va dasturiy amalga oshirish; (2) ularning individual bashorat sifatini standart mezonlar bo‘yicha baholash; (3) Tixonov regulyarizatsiyasiga asoslangan adaptiv kombinatsiya algoritmini formallashtirish; (4) taklif etilgan kombinatsiyaning samaradorligini eng yaxshi bazaviy model bilan qiyosiy tahlil qilish.

Maqolaning tuzilishi quyidagicha: 2-bo‘limda tadqiqot obyekti, bazaviy algoritmlar va robast kombinatsiyasining matematik formalizatsiyasi keltirilgan; 3-bo‘limda hisoblash eksperimenti natijalari taqdim etiladi; 4-bo‘limda natijalar muhokama qilinib, tadqiqotning cheklovlari ochiq bayon etiladi; 5-bo‘limda xulosalar umumlashtiriladi.

2. Usullar.

Tadqiqot obyekti sifatida Farg‘ona viloyati bo‘yicha yig‘ilgan, yurak-qon tomir kasalliklari bilan bog‘liq bemorlarning kunlik murojaatlari chastotasi olindi. Kirish o‘zgaruvchilari sifatida atmosfera bosimi, harorat, namlik, shamol tezligi, geomagnit faollik indeksi hamda ushbu parametrlarning sutkalik o‘zgarish tezliklari qabul qilingan - chunki klinik tajriba shuni ko‘rsatadiki, yurak-qon tomir kasalliklari holatiga ko‘pincha ko‘rsatkichlarning mutlaq qiymatidan ko‘ra ularning keskin o‘zgarishi hal qiluvchi ta’sir ko‘rsatadi [11]. Ma’lumotlar to‘plami N kunlik kuzatuvdan iborat bo‘lib, ko‘p o‘lchamli vaqt qatorini supervised o‘rganish strukturasiga keltirish uchun kengligi 7 kun bo‘lgan sirpanib boruvchi oyna texnikasi qo‘llanildi; natijada har bir kirish vektori 70 o‘lchamga ega bo‘ldi. Tanlama o‘qitish (85%) va sinov (15%) qismlariga bo‘lindi.

Bazaviy intellektual algoritmlar.

Ansambl tarkibiga kontseptual matematik apparati bilan bir-biridan farqlanuvchi to‘rtta bazaviy algoritm kiritildi.

Ko‘p qatlamli perseptron Hornik–Stinchcombe–White universal approksimatsiya teoremasiga asoslanadi [12] va statik nochiziqli bog‘liqliklarni ushlashda samarali. MLP ni o‘qitish uchun Levenberg-Marquardt usuli [13] qo‘llanildi - bu ikkinchi tartibli usul o‘rta hajmdagi tanlamalarda tez va barqaror yaqinlashishni ta’minlaydi:

$$\Delta w = (J^T J + \mu I)^{-1} J^T e \qquad (1)$$

bu yerda J - qoldiqlar Yakob matritsasi, μ - adaptiv siljish parametri, e - xatolar vektori.

Moslashuvchan neyro-noravshan xulosa tizimi [14] Takagi–Sugeno tipidagi noravshan qoidalarni besh qatlamli neyron tarmoq ko‘rinishida amalga oshiradi va hibrid algoritm bilan o‘qitiladi. ANFIS ning asosiy afzalligi - klinik interpretatsiyaga imkon beruvchi lingvistik qoidalar bazasi.

Uzoq-qisqa muddatli xotira tarmog‘i [15] unutish, kirish va chiqish darvozalari hamda Constant Error Carousel mexanizmi orqali uzoq muddatli vaqtli bog‘liqliklarni saqlaydi va

(cc) BY

gradient so‘nishi muammosini bartaraf etadi. Yacheyka holatining yangilanishi asosan chiziqli - element bo‘yicha qo‘shish orqali amalga oshiriladi, bu esa xato signalini vaqt bo‘ylab so‘ndirmasdan tashish imkonini beradi:

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \qquad (2)$$

bu yerda f_t, i_t - mos ravishda unutish va kirish darvozalarining sigmoid chiqishi, $\tilde{C}_t$ - nomzod yacheyka holati, $\odot$- element bo‘yicha ko‘paytma.

Boshqariluvchi rekurrent birlik (GRU) [16] LSTM ning soddalashtirilgan varianti bo‘lib, ikkita darvoza yordamida yashirin holatni konveks kombinatsiya ko‘rinishida yangilaydi:

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \qquad (3)$$

Bu tuzilma parametrlar sonini LSTM ga nisbatan taxminan 25% ga kamaytiradi va o‘qitish vaqtini qisqartiradi, aniqlikni esa deyarli saqlab qoladi [16]. Ikkala rekurrent model ham Adam optimizatori [17] bilan o‘qitildi, chunki u har bir parametr uchun moslashuvchan o‘rganish tezligini ta’minlaydi va shovqinli stoxastik gradient sharoitida barqaror ishlaydi. Barcha to‘rtta bazaviy modelning kirish-chiqish tuzilmasi bir xil sirpanib boruvchi oyna prinsipiga asoslanganligi ularni bir xil metamodelga integratsiyalashni soddalashtiradi.

Robast optimallashtirishga asoslangan adaptiv kombinatsiya algoritmi.

m ta bazaviy IA ning bashoratlari $\hat{y}_1$, $\hat{y}_2$, …, $\hat{y}_m$ berilgan bo‘lsin. Klassik chiziqli kombinatsiya yakuniy bashoratni $\hat{y} = \Sigma w_i \hat{y}_i$ ko‘rinishida quradi, bunda og‘irliklar $w = (w_1,\ldots,w_m)$ odatda eng kichik kvadratlar usuli bilan $Y = Xw + \varepsilon$ masalasini yechish orqali topiladi. Biroq bazaviy modellar bashoratlari o‘zaro yuqori korrelyatsiyaga ega bo‘lganda X^TX matritsasi yomon shartlangan yoki deyarli maxsus bo‘lib qoladi - bu klassik yechimni noturg‘un qiladi [8].

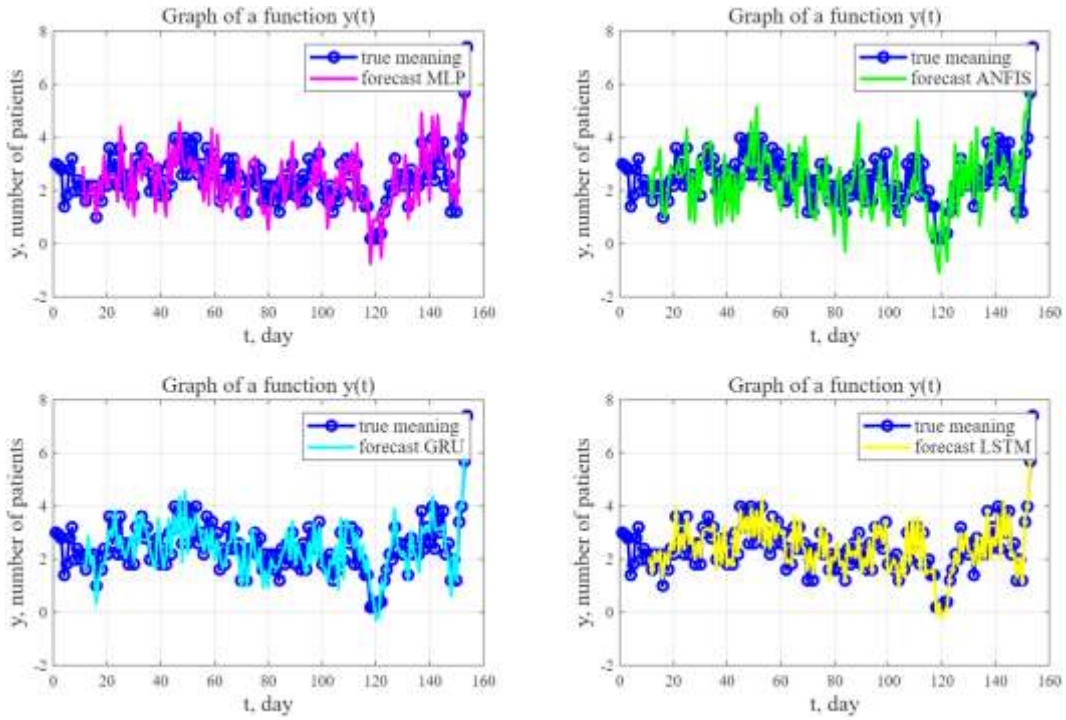
Ushbu muammoni hal etish uchun A.N. Tixonov tomonidan taklif etilgan regulyarizatsiya g‘oyasi qo‘llanildi [8, 9]: minimallashtiriladigan funksionalga qo‘shimcha jarima hadi kiritiladi:

$$J(w) = \|Y - Xw\|^2 + \lambda\|\Gamma w\|^2 \qquad (4)$$

bu yerda λ - regulyarizatsiya parametri, Γ - stabilizator operatori. Mazkur tadqiqotda stabilizator sifatida Yevklid normasining kvadrati tanlandi - bu tanlov nafaqat yechimning yagonaligini kafolatlaydi, balki shovqinli ma’lumotlar sharoitida ansambl og‘irliklarining barqarorligini sezilarli oshiradi [10]. Optimal yechim analitik ko‘rinishga ega:

$$w^* = (X^TX + \lambda I)^{-1}X^TY \qquad (5)$$

Regulyarizatsiya parametri λ iterativ algoritm asosida aniqlanadi; iteratsiyani to‘xtatish sharti sifatida w qiymatlarining ketma-ket iteratsiyalar orasidagi barqarorlashishi qabul qilindi. Amalga oshirishda A2 algoritmi statistika yetarli to‘plangan vaqt oynasidan boshlab ishga tushiriladi va MATLAB Optimization Toolbox vositalari yordamida hisoblandi.



1-rasm. Bemorlarning murojaati chastotasining bashorat va haqiqiy qiymatlari.

(cc) BY

3. Natijalar

To‘rtta bazaviy algoritmning sinov tanlamasidagi yakuniy ko‘rsatkichlari 1-jadvalda keltirilgan.

1-jadval.

4 IA modellarni natijalari.

Model	MAE	RMSE	O‘qitish vaqti, s	Asosiy afzalligi
ANFIS	0.87	0.9105	2	Eng tezkor
MLP	0.66	0.7304	43	Tezkor, universal
GRU	0.41	0.5102	200	Kamroq parametr
LSTM	0.26	0.2906	250	Eng yuqori aniqlik

1-jadvaldan ko‘rinib turibdiki, aniqlik bo‘yicha modellar quyidagi tartibda joylashadi: LSTM > GRU > MLP > ANFIS, aniqlik va o‘qitish vaqti orasida esa teskari proporsional bog‘liqlik kuzatiladi – eng aniq model eng ko‘p vaqt talab qiladi, eng tezkor model esa eng past aniqlikka ega. Bundan tashqari, har bir modelning 170ató strukturasi vaqt oynalari bo‘yicha qisman korrelyatsiyalanmagan bo‘lib chiqdi: ayrim oraliqlarda LSTM eng aniq 170ató170a bersa, boshqa oraliqlarda GRU yoki MLP ustunlik qildi. Krogh va Vedelsby ansambl 170ató dekompozitsiyasi nazariyasiga ko‘ra [18], ansambl bashoratining kvadratik xatosi bazaviy modellar xatolarining vaznli o‘rtachasidan bazaviy modellar bashoratlari orasidagi tafovutga proporsional miqdorga kam bo‘ladi – ya’ni ansambl xatosi bazaviy modellar o‘rtacha xatosidan oshmaydi. Aynan shu tafovutning mavjudligi ansambllashtirish g‘oyasining nazariy asosini tashkil etadi.

2-jadval.

RMSE ko‘rsatkichlar.

Ko‘rsatkich	MLP	ANFIS	GRU	LSTM
RMSE	0.6667	0.8729	0.4148	0.2600

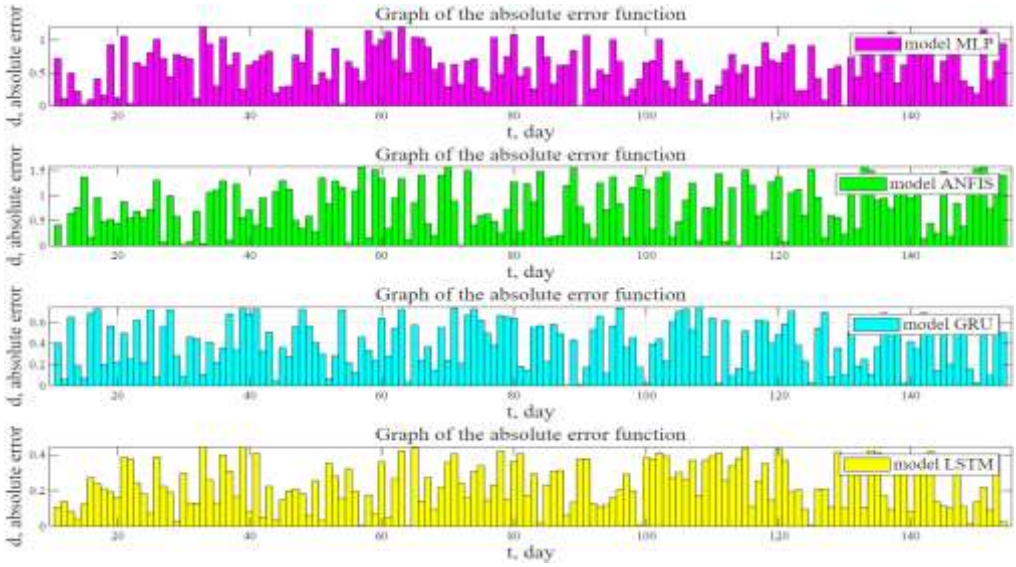
Robast kombinatsiyasi natijalari.

A2 algoritmi bo‘yicha aniqlangan og‘irliklar asosida shakllantirilgan yakuniy bashorat bazaviy modellar bilan solishtirildi. Robast kombinatsiyaning RMSE ko‘rsatkichi 0.1985 ni tashkil etdi – bu eng yaxshi bazaviy model ko‘rsatkichidan taxminan 23,7% past.

3-jadval.

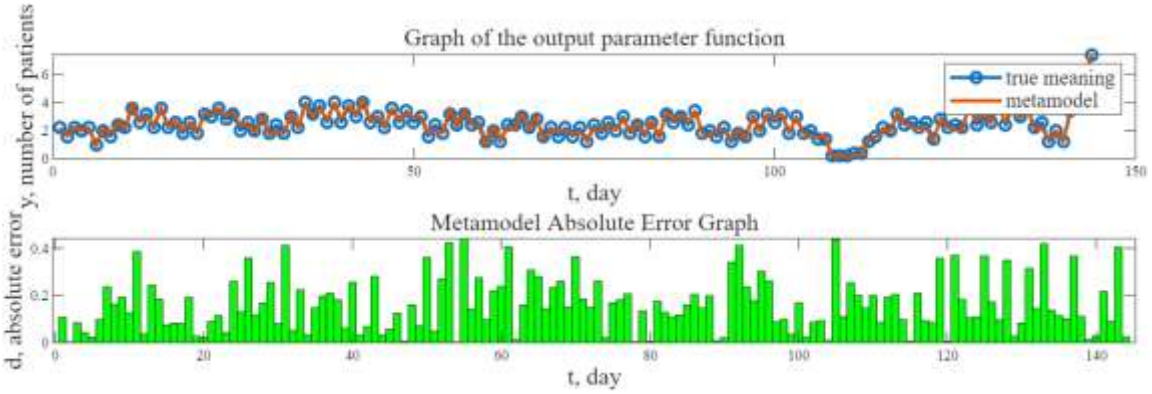
Eng yaxshi IA model va ansabl modellar solishtirmasi.

Usul	RMSE
Eng yaxshi bazaviy model (LSTM)	0.2600
Vaznli o‘rtacha (A1)	~0.23–0.25*
Tixonov robast kombinatsiyasi (A2)	0.1985



2-rasm. Vaqt oynalari bo‘yicha bemorlarning murojaati chastotasini bashoratlashda IA larning mutlaq xatolari diagrammasi.

(cc) BY *dtai.uz*



3-rasm. IA ansambli orqali bemorlarning murojaati chastotasini bashoratlash natijalari.

4. Muhokama.

Olingan natijalar robast kombinatsiyaning bazaviy modellar orasidagi yuqori korrelyatsiya sharoitida ham turg‘un va shovqinga bardoshli og‘irliklarni aniqlash imkonini berishini tasdiqlaydi. Bu natija Bertsimas va Boussioux [10] tomonidan boshqa sohalarda olingan xulosalarga mos keladi.

Bu natijaning amaliy ahamiyati shundaki, robast kombinatsiya yakka modellarning kuchli tomonlarini birlashtirib, ularning zaif tomonlarini vaqt bo‘yicha o‘zaro qopladi: LSTM ning uzoq muddatli xotirasi, GRU ning tezkorligi, MLP ning statik nochiziqlilikni yaxshi ushlashi va ANFIS ning silliqlash xususiyati bir vaqtning o‘zida hisobga olindi. Bunda Tixonov regulyarizatsiyasining roli ayniqsa muhim: modellar bashoratlari orasidagi yuqori korrelyatsiya klassik eng kichik kvadratlar yechimini beqaror qiladigan bo‘lsa-da, regulyarizatsiya hadi $\lambda\|w\|^2$ og‘irliklarning haddan tashqari katta yoki manfiy qiymatlarga chiqib ketishining oldini oldi va shu orqali ekstremal kuzatuvlar sharoitida ham bashoratning barqarorligini ta’minladi.

Shu bilan birga, tadqiqotning bir qator cheklovlarini ochiq tan olish lozim. Birinchidan, natijalar bitta ma’lumotlar to‘plamida bitta hisoblash eksperimenti asosida olingan bo‘lib, statistik ishonchlilik oralig‘i yoki bir necha marta takrorlangan sinovlar bo‘yicha dispersiya keltirilmagan; bu - natijalarning umumlashtirilishini cheklovchi muhim omil va kelgusi nashrlarda albatta bartaraf etilishi lozim. Ikkinchidan, taqqoslash faqat klassik neyron tarmoq va neyro-noravshan arxitekturalar bilan olib borilgan; zamonaviy transformer-asosidagi bashoratlash modellari bilan qiyosiy tahlil ushbu maqola doirasidan tashqarida qoldi va kelgusi tadqiqotlar mavzusi bo‘lib qolmoqda - bunday modellar katta hajmdagi ma’lumotlarda ustunlik ko‘rsatishi mumkin, biroq mazkur tadqiqotdagi nisbatan cheklangan hajmdagi tibbiy-meteorologik tanlamada ularning afzalligi oldindan aniq emas. Uchinchidan, taklif etilgan metodologiyaning universalligi haqidagi da’vo faqat bitta tibbiy-meteorologik obyekt misolida tekshirilgan bo‘lib, boshqa sohalarga tatbiqi qo‘shimcha empirik tasdiqni talab qiladi. To‘rtinchidan, vaznli o‘rtacha (A1) algoritmi bo‘yicha aniq raqamli natija ushbu maqolada taqqoslash maqsadida taxminiy ko‘rsatilgan bo‘lib, uni alohida hisoblash eksperimentida aniqlashtirish zarur.

5. Xulosa.

Maqolada sust shakillangan jarayonlarni bashoratlash uchun to‘rtta kontseptual jihatdan farqli bazaviy intellektual algoritm qurilgan va ularning bashoratlarini robast optimallashtirish nazariyasiga asoslangan holda birlashtiruvchi adaptiv kombinatsiya algoritmi taklif etilgan. yurak-qon tomir kasalliklari bilan bog‘liq bemorlar murojaati chastotasini bashoratlash misolida o‘tkazilgan hisoblash eksperimenti robast kombinatsiyaning eng yaxshi bazaviy modeldan yuqori aniqlikka erishishini ko‘rsatdi. Kelgusi tadqiqotlar statistik ishonchlilikni bir necha marta takrorlangan sinovlar orqali tasdiqlashga, transformer-asosidagi zamonaviy modellarni taqqoslashga hamda metodologiyani boshqa sohalarga kengaytirishga qaratilishi maqsadga muvofiq.

(cc) BY

Foydalanilgan adabiyotlar

1. Lim B., Zohren S. Time-series forecasting with deep learning: a survey // Philosophical Transactions of the Royal Society A. – 2021. – Vol. 379. – Article 20200209.
2. Yang H., Bukkapatnam S.T.S. Time series forecasting for nonlinear and non-stationary processes // IIE Transactions. – 2015. – Vol. 47, No. 10. – P. 1053–1071.
3. World Health Organization. Cardiovascular diseases (CVDs): Fact sheet. – Geneva: WHO, 2024.
4. Hewamalage H., Bergmeir C., Bandara K. Recurrent Neural Networks for Time Series Forecasting // International Journal of Forecasting. – 2021. – Vol. 37, No. 1. – P. 388–427.
5. Gama J., Žliobaitė I., Bifet A. et al. A survey on concept drift adaptation // ACM Computing Surveys. – 2014. – Vol. 46, No. 4. – Article 44.
6. Dietterich T.G. Ensemble methods in machine learning // Multiple Classifier Systems, LNCS vol. 1857. – Springer, 2000. – P. 1–15.
7. Wolpert D.H. Stacked generalization // Neural Networks. – 1992. – Vol. 5, No. 2. – P. 241–259.
8. Тихонов А.Н. Об устойчивости обратных задач // Доклады Академии наук СССР. – 1943. – Т. 39, № 5. – С. 195–198.
9. Tikhonov A.N., Arsenin V.Y. Solutions of Ill-Posed Problems. – Washington, D.C.: V.H. Winston & Sons, 1977.
10. Bertsimas D., Boussioux L. Ensemble Modeling for Time Series Forecasting: an Adaptive Robust Optimization Approach // arXiv:2304.04308. – 2023.
11. Wang H., Liu Y., Zhou B. et al. The relationship between meteorological parameters and daily emergency department visits for cardiovascular diseases // Frontiers in Public Health. – 2022. – Vol. 10. – Article 853192.
12. Hornik K., Stinchcombe M., White H. Multilayer feedforward networks are universal approximators // Neural Networks. – 1989. – Vol. 2, No. 5. – P. 359–366.
13. Hagan M.T., Menhaj M.B. Training feedforward networks with the Marquardt algorithm // IEEE Transactions on Neural Networks. – 1994. – Vol. 5, No. 6. – P. 989–993.
14. Jang J.-S.R. ANFIS: Adaptive-network-based fuzzy inference system // IEEE Transactions on Systems, Man, and Cybernetics. – 1993. – Vol. 23, No. 3. – P. 665–685.
15. Hochreiter S., Schmidhuber J. Long short-term memory // Neural Computation. – 1997. – Vol. 9, No. 8. – P. 1735–1780.
16. Cho K., van Merriënboer B., Gulcehre C. et al. Learning phrase representations using RNN encoder-decoder // Proceedings of EMNLP 2014. – 2014. – P. 1724–1734.
17. Kingma D.P., Ba J. Adam: A method for stochastic optimization // Proceedings of ICLR 2015. – 2015.
18. Krogh A., Vedelsby J. Neural network ensembles, cross validation, and active learning // Advances in Neural Information Processing Systems. – 1995. – Vol. 7. – P. 231–238.

(cc) BY