

3. A. Mohamed and N. Jaitly, 2013, “Hybrid speech recognition with deep bidirectional LSTM,”, IEEE, 2013, pp. 273–278.

4. Ниёзматова, Н. А., Маматов, Н. С., Отахорова, Б. И., Бобоев, Л. Б., & Самижонов, А. Н. Матнларни таснифлашда информатив белгилар мажмуасини аниқлаш усуллари.

МАЪЛУМОТЛАРНИ ИНТЕЛЛЕКТУАЛ ТАҲЛИЛЛАШ УЧУН БАШОРАТЛАШ ЁНДАШУВЛАРИ

Бабомурадов Озод Жураевич

т.ф.д.(DS), проф. Жиззах шаҳридаги ҚФУ филиали

bobomuradov@gmail.com

Хайдаров Озоджон Асламкулович

Малака ошириш йўналиши раҳбари

“Creative Associates International” Ўзбекистондаги ваколатхонаси

haydarov@gmail.com

Аннотация: Ушбу мақола бошқарув тизимларида сифатли қарорлар қабул қилиш учун маълумотлар асосида башорат моделларини ишлаб чиқиш заруратига бағишланган. Маълумотларни таҳлил қилиш ва башоратлаш жараёнида вақтли қаторлар асосидаги усуллар кўрсатилган. Вақтли қаторлар анализи ва бошқарув стратегиясини шакллантиришдаги аҳамияти, аниқлик ва самарадорликка таъсири ҳамда натижаларнинг ишончлилигини ошириш учун зарур бўлган математик модель ва алгоритмлар таҳлили келтирилган.

Калит сўзлар: NNS, MASE, RF, TREE Бағлар, SMAPE нейрон тармоғи, XG boost, PCA, FFORMS – белгилар, SDAE, SAE, MLP, ARIMA, PropHet, BSE, BC, Ada Boost, NYSE, NASDAQ, Exca Trees, GSO, SYM, KNN, LSTM, GRU, NSE, FinBERT, ELM, GB, KNN, DT, SARIMA, тасодифий ўрмон, SVM, чизикли регрессия.

ПОДХОДЫ К ПРОГНОЗИРОВАНИЮ ДЛЯ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ

Аннотация: Данная статья посвящена необходимости разработки моделей прогнозирования на основе данных для принятия качественных решений в управленческих системах. Рассмотрены методы, основанные на временных рядах, в процессе анализа данных и прогнозирования. Приводятся анализ временных рядов, влияние на формирование управленческой стратегии, а также математические модели и алгоритмы, необходимые для повышения точности и эффективности, а также улучшения надежности результатов.

Ключевые слова: NNS, MASE, RF (Случайный лес), TREE (Деревья), SMAPE, Нейронная сеть, XGBoost, PCA (Анализ главных компонент), FFORMS – признаки, SDAE, SAE, MLP (Многослойный перцептрон), ARIMA, Prophet, BSE, BC, AdaBoost, NYSE, NASDAQ, Extra Trees, GSO, SYM, KNN (Метод k-ближайших соседей), LSTM, GRU, NSE, FinBERT, ELM (Экстремальная машина обучения), GB (Градиентный бустинг), DT (Дерево решений), SARIMA, Случайный лес, SVM (Метод опорных векторов), Линейная регрессия.

PREDICTIVE APPROACHES FOR INTELLIGENT DATA ANALYSIS

Abstract: This article is dedicated to the necessity of developing predictive models based on data for making quality decisions in management systems. Methods based on time series in the

process of data analysis and forecasting are presented. The importance of time series analysis in shaping management strategy, its impact on accuracy and efficiency, as well as the analysis of mathematical models and algorithms necessary to enhance the reliability of results, are discussed.

Keywords: NNS, MASE, RF (Random Forest), TREE, SMAPE, Neural Network, XGBoost, PCA (Principal Component Analysis), FFORMS – features, SDAE, SAE, MLP (Multilayer Perceptron), ARIMA, Prophet, BSE, BC, AdaBoost, NYSE, NASDAQ, Extra Trees, GSO, SYM, KNN (K-Nearest Neighbors), LSTM, GRU, NSE, FinBERT, ELM (Extreme Learning Machine), GB (Gradient Boosting), DT (Decision Tree), SARIMA, Random Forest, SVM (Support Vector Machine), Linear Regression.

Кириш. Маълумки бошқарув тизимларида сифатли қарор қабул қилишга қўмаклашувчи қарор муқобиллигини ишлаб чиқиш учун маълумотларга ишлов беришнинг самарали модел ва алгоритмларининг бўлишлиги талаб этилади.

Маълумотлар массиви гипер ўсиши кузатилаётган бир вақтда унга ишлов беришнинг воситалари турлари ҳам ортиб бормоқда. Бошқарув стратегиясининг тўғри шаклланиши бошқарув самарадорлигини оширишдаги асосий омилларидан бири ҳисобланади. Ушбу стратегия аниқлиги, таъсир омили сифатида кучини ҳолатларни тўғри башоратлай билиш белгилаб беради [1,2].

Башорат масаласи бирор ҳолат ёки ходисани вақт оралиғидаги ўзгариш динамикасини тадқиқ қилиш орқали маълум қонуниятни ҳосил қилган ҳолда навбатдаги қадам ёки босқич (итерация) даги ўзгаришларни (яширин) қонуниятларни юзага чиқаришдан иборат бўлади. Бунинг учун қўйилган боғлиқликлар ва турли параметрликларни амалга ошириш талаб этилади.

Қиймат кўрсаткичлар ўзида бир ўлчамли векторлар ёки сон қиймат birlikлари массивини мужассамлаштиради. Массивда амалга ошириладиган ҳар бир итерация маълум бир вақт оралиғига боғланган бўлади, буни ахборот рақамли ишлов бериш соҳасида вақтлик қатор сифатида тавсифланади[3,4].

Турли соҳаларда вақтлик қатор турига таъсир кучини ва йўналтирмасини белгилашга хизмат қилиши мумкин. Бунда вақт оралиғи жараён ўзгаришларига турли таъсири ёки ҳолатни белгилаш тарзда иштироки таминланади. Масалан, молиявий вақтлик қатор қўлланилганда таннарх кўрсаткичлари ва жараёнлари бази кўринишдаги тебранишларни аниқлашга қаратилади, бу молиявий кўрсаткичнинг башорат зўнасидаги ўзини тутишидаги зарурий сигналлар, қонуниятлар ҳамда белгиларни интерпретациясида ҳалақитларни келтириб чиқариши мумкин. Бу эса ўз навбатида жараёни интеллектуал таҳлили самарадорлигини пасайишига, трендни нотоғри талқини ёки натижани олишдаги кечикишга олиб келади. Жараёндаги интеллектуал таҳлил самарадорлигини пасайиши, вақтлик қатор асосида башорат модели ишчи ҳолатини қониқарсиз кўринишга келтиради, трендни нотўғри талқини эса башорат йўналишини бутунлай бошқа томонга бурилиб кетишини кузатиш мумкин бўлади. Интелектни олишдаги кечикиш башоратнинг тезкорлик қияматига салбий таъсир ўтказади. Бу ҳолат башоратнинг ўзининг заруратлилиги масаласига шубҳа ҳосил қилиши мумкин [5].

Маълумотлар таҳлили ва башоратлар жараёни вақтлик қаторларга дастлабки ишлов бериш механизмини такомиллаштириш орқали эришилади. Бунинг учун бир қатор йўналишлар борки ушбу дастлабки ишлов бериш моделларини ўринли қўллаш вақтлик қатор маълумотлардаги нуқсонларга барҳам бериш, маълумотларни текислигини имконини беради. Тозаликни вақтлик қаторлар нейрон бирламчи ёки унинг бошқа математик моделлар билан таянган ҳолда амалга оширилади [6].

Вақтли қаторларни башоратлаш масалаларида вақтга боғланган ҳолда масалани ҳал этиш ёндашуви қараб ўтилади. Бунда асосий масала сифатида дискретлаш қаралиб, $x(k)$ харажат кўрсаткичларининг узлуксиз функциясини $x(t)$ вақт қатори кўринишида ифодалашда дискретлаш жараёни назарда тутилади.

$$x(t) = x(t) \sum_{k=0}^{\infty} (\delta) x(t - kT) = \sum_{k=0}^{\infty} x(kT) \delta(t - kT) \quad (1)$$

Фуре конвертациясини қўллаш орқали

$$s(\omega) = \int_0^{\infty} x(t) e^{-j\omega t} dt \quad (2)$$

вақт сериясининг интеграл ости қийматини алмаштиргандан сўнг шаклнинг $x(kT)$ шаклида (1), вақт сериясининг спектрал функцияси куйидаги тарбида ҳосил қилинади

$$S_T(\omega) = \sum_{k=0}^{\infty} x(kT) \int_0^{\infty} e^{-j\omega t} \delta(t - kT) dt = \sum_{k=0}^{\infty} s(kT) e^{-j\omega kT} \quad (3)$$

Комплекс частота $j\omega$ ўрнига $p = \delta + j\omega$ комплекс ўзгарувчини қўйиб, Лаплас кўринишидаги вақт қаторининг ифодасини ҳосил қилиб олинади.

$$S_T(p) = \sum_{k=0}^{\infty} s(kT) e^{-pkT} \quad (4)$$

Лаплас вақт сериясининг $S_T(p)$ тасвирини $Z = e^{pT}$ ифодаси орқали P ўзгарувчиси билан боғланган янги Z ўзгарувчисига ўтиш орқали соддалаштиришни амалга ошириш мумкин.

Лаплас бўйича Z қийматини вақт сериясининг тасвирига алмаштиришда Z -текислигидаги вақт сериясининг тасвири шаклда олинади.

$$S_T(z) = \sum_{k=0}^{\infty} s(kT) z^{-k} \quad (5)$$

Z -ўзгартириш вақт сериясини қайта ишлаш алгоритмларини филтрлаш узатиш функциялари ва узатиш функцияларини синтез қилиш учун ишлатилади:

$$k(z) = \sum_{k=0}^{\infty} g(kT) z^{-k} \quad (6)$$

Бу ерда $g(kT)$ - ишлов бериш алгоритмининг тўрт портли тармоғида импульсли акслантиришни ифодалайди.

Асл нусха тескари Z -ўзгартириш орқали олинади.

$$g(kT) = \frac{1}{2\pi j} \oint_{|z|=1} k(z) z^{(k-1)} dz \quad (7)$$

$x(k)$ сигнал намуналарининг k кетма-кетлигини N тартибли намунадаги Фурье конвертацияси шакли олинади.

$$S(e^{j\omega}) = \sum_{k=0}^{N-1} x(k) e^{-j\omega kT} \quad (8)$$

Бундай ҳолда, ω_l 0 дан $2\pi F_{\partial}$ гача ўзгарганда Z -текислигининг бирлик доираси бўйлаб бир текис тақсимланган L нукталарида Дискрет Фурье ўзгартиришнинг қийматларини ҳисоблаш ифода бўйича амалга оширилади.

$$S\left(e^{j\frac{2\pi F_{\partial}}{L} \cdot l}\right) = \sum_{l=0}^{L-1} \sum_{k=0}^{N-1} x(k) e^{-j2\pi \frac{F_{\partial}}{L} l k T}, \quad L < N \text{ да} \quad (9)$$

$L = N$ ҳолатида (9) ифодаси бундай шаклга эга

$$S\left(e^{j\frac{2\pi F_{\partial}}{N} \cdot l}\right) = \sum_{l=0}^{N-1} \sum_{k=0}^{N-1} x(k) e^{-j2\pi \frac{F_{\partial}}{N} l k T} \quad (10)$$

Юқоридаги Дискрет Фурье ўзгартириши ифода, $W_N = e^{-j\frac{2\pi}{N}}$ белгисини қабул қилганда, яъни $S(l) = S(j \cdot \Delta\omega \cdot l) = S(j \cdot \omega_l)$ бўлса, куйидагича ифодаланади

$$S(1) = \sum_{k=0}^{N-1} x(k) w_N^{1k}, \text{ if } L = N, \text{ else } 0 \leq l \leq N - 1 \quad (11)$$

Тескари дискрет Фурье ўтказиш куйидаги ифода ёрдамида амалга оширилади.

$$x(k) = \frac{1}{N} \sum_{l=0}^{N-1} S(l) w_N^{-lk} \quad (12)$$

Тўлқинлар стационар бўлмаган сигналларни таҳлил қилиш ва қайта ишлашнинг қулай математик усулларида бири бўлиб, улар харажатлар кўрсаткичлари ўзгаришининг узлуксиз функциясининг дискрет тасвирини ўз ичига олади. Тўлқинли таҳлил назариясига кўра, ҳар қандай функция маълум даражада $i = n$ кўринишдаги қаторга кенгайтирилиши мумкин

$$s(k) = \sum_k C_{n,k} \varphi_{n,k} + \sum_k \sum_{i=1}^n d_{i,k} \Psi_{i,k} \quad (13)$$

бу ерда $\varphi_{n,k}$ ва $\psi_{i,k}$, масштаблаш функцияси ва тўлқинли функцияси кўринишга эга.

$$\Psi(t) = \sqrt{2} \sum_1 g_1 \varphi(2t - k), \quad (14)$$

бу ерда $l = 0, 1, \dots, l_0 = 2m - 1$, m – тўлқинлар тартиби.

Масштаблаш функциясининг коэффициентлари (масштаблаш вектори ёки масштаблаш филтри) h_l қайта ишлаш филтрларининг вақтинчалик характеристикаларини ифодалаши мумкин.

Шундай қилиб, сигналларнинг тўлқинли парчаланишига асосланган ишлов бериш моделини яратиш усули автоматлаштирилган мониторинг тизимларида дастлабки ишлов бериш алгоритмларини яратиш учун истиқболли йўналишлардан ҳисобланади.

Аксарият тизимларда машинали ўқитишнинг регрессия таҳлилини қўллаш орқали маълумотлар таҳлили амалга оширилади. Бу ерда жараён маълум бўлиши маълумотларга асосланган параметрлар функцияларни кўришда қўлланилади. Чизиқли регрессия параметрлар таҳлилида камчиликларга йўл қўйишни бошласа унинг турли математик аппаратлар билан аралашмаси гибридлашувидан фойдаланамиз.

Бунинг учун:

- башорат учун регрессия модели ёки унинг модификацияси таинланади;
- регрессия моделлари асосида башоратга таъсир этувчи белгиларни танлаш;
- башоратлашнинг чуқур ўқитишга асосланган моделини ишлаб чиқиш.

Башорат учун зарур бўладиган вақтли қаторларни олиш, топлаш ва сақлаш жуда кўп амалга ошириш йўллари мавжуд. Молиявий кўрсаткичлар маълум таъсирлар ва талаб бўладиган вақт оралиқларида турлича бўлиши мумкин. Тебранишларга тасирлар ҳам ҳудди шу каби ўзининг динамикасига эга бўлади.

Бошқа иқтисодиёт соҳалари ҳам ўзининг мос равишда таъсир омилларини касб этган ҳолда намоён бўлади. Бу ҳолат вақтли қаторни ҳосил қилишда реал вақтга асосланса таҳлиллаш (ишлов бериш) механизмида ҳам шу тарзда ўзгариш қилишни талаб этади, яни мазкур ўзгарувчига динамик ўзгарувчи мақоми берилади. Мазкур ҳолатда маълумот ўзгариши баробарида доимий адаптив тарзда маълумотлар базасига ёзилади серверга сақланиб боради. Бу ҳолат тезкор ишлов бериш жараёни бўлиб, кўпроқ биржадаги кодировка ўзгаришларини қайд этишда кузатилиши мумкин.

Кўп ҳолларда вақтли қаторларга ишлов беришга қулай математик аппарат сифатида спектрал таҳлил назарияси, Фурье Лаплас ва бошқа кўринишдаги ўзгаришлардан фойдаланилади. Тадқиқот жараёнида мазкур ёндашувлар ўринли ҳолатларда фойдаланиши назарда тутилган.

Адабиётлар рўйхати

1. Bekmuratov T.F., Mukhamedieva D.T., Bobomuradov O.J. Model prediction of yielded with fuzzy initial conditions// Ninth International Conference on Application of Fuzzy Systems and Soft Computing: ICAFS-2010. – August 26-27, 2010. b-Quadrat Verlag - Prague, Czech Republic, 2010. – p.321-328.

2. Бекмуратов Т.Ф., Мухаммадиева Д.Т., Бобомуратов О.Ж., Акбаралиев Б.Б. Модели принятия слабоструктурированных решений// Труды Восьмой Международной симпозиум “Интеллектуальные системы”. – г.Нижний Новгород, Россия, 2008. – С. 437-441.

3. Бобомуратов О.Ж., Хамраев А.Ш. Ноаниқлик шароитида прогнозлаш масаласи//Амалий математика ва информацион технологияларнинг долзарб муаммолари. ал-Хоразмий 2009, 2-тўплам, 18-21 сентябрь 2009.-Б.76-77.

4. Sasan Barak, Mahdi Nasiri, Mehrdad Rostamzadeh (2019) Time series model selection with a meta-learning approach; evidence from a pool of forecasting algorithms// https://www.semanticscholar.org/paper/Time-series-model-selection-with-a-meta-learning-a-Barak-Nasiri/e52508600b806171bd05bcb11dc10ad16c9461bf?utm_source=direct_link

5. Thiyanga S. Talagala, Rob J. Hyndman, George Athanasopoulos Meta-learning how to forecast time series //Journal of Forecasting. 2023;42:1476–1501./ © 2023 John Wiley & Sons Ltd. DOI: 10.1002/for.2963

6. Bandara, K., Hyndman, R. J., & Bergmeir, C. (2021). MSTL: A seasonal-trend decomposition algorithm for time series with multiple seasonal patterns. arXiv preprint 2107.13462. <https://arxiv.org/abs/2107.13462>

MATNLARNI TEMATIK SINFLASHTIRISH VA SAMARADORLIGINI TAHLIL QILISH

Muhammediyeva Dilnoz Tulkunovna

“Raqamli texnologiyalar va sun’iy intellekt” kafedrası professori, Toshkent irrigatsiya va qishloq xo‘jaligini mexanizatsiyalash muhandislari instituti Milliy tadqiqot universiteti

Mamatov Abduvali Abduvohid o‘g‘li

Namangan davlat universiteti, 1-bosqich tayanch doktoranti, Raqamli texnologiyalar va sun’iy intellekt kafedrası assistenti, Toshkent irrigatsiya va qishloq xo‘jaligini mexanizatsiyalash muhandislari instituti Milliy tadqiqot universiteti

Turg‘unova Nafisaxon Maxammadjon qizi

Raqamli texnologiyalar va sun’iy intellekt kafedrası assistenti, Toshkent irrigatsiya va qishloq xo‘jaligini mexanizatsiyalash muhandislari instituti Milliy tadqiqot universiteti

Anotatsiya: Ushbu maqolada turli mashinaviy o‘rganish modellarining tasniflash samaradorligi tahlil qilinadi. Matnlarni raqamli formatga o‘tkazish uchun TF-IDF (Term Frequency-Inverse Document Frequency) usuli qo‘llaniladi. Sinov uchun Logistic Regression, Ridge Classifier, k-Nearest Neighbors (kNN), Random Forest, Linear SVC, Stochastic Gradient Descent (SGD), Nearest Centroid va Complement Naïve Bayes kabi klassifikatorlar tanlangan. Modellar aniqlik darajasi, o‘qitish vaqti va test vaqti bo‘yicha taqqoslanadi. Xatolik matritsalar yordamida har bir modelning turli kategoriyalarni tasniflashdagi natijalari vizualizatsiya qilinadi. Natijalarga asoslangan holda eng samarali model tanlanadi va uning afzalliklari tahlil qilinadi.

Kalit so‘zlar: Matn klassifikatsiyasi, TF-IDF, mashinaviy o‘rganish, tasniflash, xatolik matritsa, model samaradorligi, Stochastic Gradient Descent, Random Forest, Linear SVC.

ТЕМАТИЧЕСКАЯ КЛАССИФИКАЦИЯ ТЕКСТОВ И АНАЛИЗ ИХ ЭФФЕКТИВНОСТИ

Аннотация: В данной статье анализируется эффективность различных моделей машинного обучения для классификации текстов. Для преобразования текстов в цифровой формат используется метод TF-IDF (Term Frequency-Inverse Document Frequency). В