



2024-2025 წლების საქართველოს მაკროეკონომიკური მაჩვენებლების პროგნოზირება, მანქანური სწავლების ჰიბრიდული მოდელის გამოყენებით

ზურაბ ტუსკია

ინფორმატიკისა და მართვის სისტემების ფაკულტეტი, საქართველოს ტექნიკური უნივერსიტეტი, კოსტავას ქ. 77, თბილისი, 0160, საქართველო

ანოტაცია

ეკონომიკური პოლიტიკის ანალიზი და შეფასება, როგორც რეალურ, ისე სიმულაციურ გარემოში, დიდწილად ეფუძნება მაკროეკონომიკური მდგომარეობის სანდო და ზუსტ პროგნოზირებას. აღნიშნული ამოცანის გადასაჭრელად ნაშრომში განხილულია ღრმა ნეირონული ქსელების არქიტექტურული შესაძლებლობები და შემოთავაზებულია ჰიბრიდული მოდელი, რომელიც აერთიანებს Long Short-Term Memory (LSTM) კომპონენტს, attention-მექანიზმს და nudging-ზე დაფუძნებულ კორექციულ მიდგომას.

მაკროეკონომიკური ინდიკატორების პროგნოზირების ხარისხის გასაუმჯობესებლად ჩამოყალიბებულია ადაპტირებადი ალგორითმული სისტემა, რომელიც ჰიბრიდული ნეირონული ქსელის ფარგლებში აერთიანებს დროითი მწკრივების დამუშავების, მნიშვნელოვანი ცვლადების იდენტიფიცირებისა და პროგნოზის კორექციის მექანიზმებს. აღნიშნული მიდგომა პროგნოზის შედეგებს ეტაპობრივად აუმჯობესებს მონაცემთა ასიმბალანსის პრინციპზე დაფუძნებული nudging-კორექციის გამოყენებით.

ექსპერიმენტული შეფასებისათვის გამოყენებულ იქნა საქართველოს რეალური მაკროეკონომიკური მონაცემები. მიღებულმა შედეგებმა აჩვენა, რომ შემოთავაზებული სისტემა უზრუნველყოფს პროგნოზირების მაღალ სიზუსტეს და ინარჩუნებს მდგრადობას მოდელის გარე პარამეტრების ცვლილების პირობებშიც.

საკვანძო სიტყვები: ჰიბრიდული მოდელი; მონაცემთა ასიმბალანსი; attention-მექანიზმი; nudging-კორექცია; ღრმა სწავლება; ნეირონული ქსელები; მაკროეკონომიკური ინდიკატორები.

1. შესავალი

მაკროეკონომიკური ინდიკატორების პროგნოზირების სისტემის აგებისას მიზანშეწონილია გამოყენებულ იქნეს მრავალკომპონენტური მოდელური არქიტექტურა, რომელიც აერთიანებს როგორც კლასიკურ ეკონომეტრიკულ მიდგომებს, ისე თანამედროვე მანქანური სწავლებისა და ღრმა ნეირონული ქსელების მეთოდებს. ასეთი მიდგომა განსაკუთრებით აქტუალურია იმ შემთხვევაში, როდესაც პროგნოზირებას ექვემდებარება ისეთი ცვლადები, როგორცაა მთლიანი შიდა პროდუქტი, ინფლაცია, უმუშევრობა, სახელმწიფო შემოსავლები, ფულადი მასა, ინვესტიციები და სხვა მაკროეკონომიკური მაჩვენებლები.

პროგნოზირების პროცესში მნიშვნელოვანია არა მხოლოდ ისტორიული ტრენდებისა და განმეორებადი დინამიკის დაფიქსირება, არამედ იმ სტრუქტურული ურთიერთკავშირების იდენტიფიცირებაც, რომლებიც ეკონომიკურ ცვლადებს შორის არსებობს. სტრუქტურული მიდგომა საშუალებას იძლევა გაანალიზდეს, როგორ შეიძლება აისახოს კონკრეტული ეკონომიკური პოლიტიკა, საბიუჯეტო გადაწყვეტილება, საინვესტიციო ცვლილება ან სხვა ტიპის გარე ზემოქმედება პროგნოზირებად მაჩვენებლებზე. ამასთან, არასტრუქტურული მოდელები ეფექტიანია მაშინ, როდესაც პრიორიტეტულია მონაცემებზე დაფუძნებული პროგნოზი და არა წინასწარ განსაზღვრული თეორიული კავშირის ფორმალიზება.

მაკროეკონომიკური პროგნოზირების თანამედროვე ლიტერატურაში პროგნოზირების მეთოდები ხშირად იყოფა ორ ძირითად მიმართულებად: სტრუქტურულ და არასტრუქტურულ მიდგომებად. არასტრუქტურული პროგნოზირება ფართოდ გამოიყენება ემპირიულ კვლევებში, მათ შორის ბიუჯეტური და ფინანსური მაჩვენებლების წინასწარი შეფასებისას [13], [14], [15], [16]. აღნიშნული მიდგომა განსაკუთრებით პრაქტიკულია მაშინ, როდესაც მონაცემთა დინამიკა რთულია, ცვლადებს შორის კავშირები კი დროში ცვალებადია.

პირველ ეტაპზე შეიძლება გამოყენებულ იქნეს კლასიკური სტატისტიკური მოდელები, მათ შორის ARIMA, SARIMA და ექსპონენციალური გასწორების მეთოდები. ეს მოდელები ეფექტიანია სტაბილური დროითი მწკრივებისათვის, განსაკუთრებით მაშინ, როდესაც მონაცემებში შეინიშნება სეზონურობა, პერიოდულობა ან შედარებით მარტივი ტენდენციური სტრუქტურა. მათი უპირატესობა მდგომარეობს ინტერპრეტაციის სიმარტივესა და ეკონომეტრიკულ გამჭვირვალობაში.

მეორე მიმართულებას წარმოადგენს მანქანური სწავლების მოდელები, რომლებიც უკეთ უმკლავდებიან მრავალცვლადიან მონაცემთა სტრუქტურას. Random Forest შეიძლება გამოყენებულ იქნეს რთული ურთიერთკავშირების გამოსავლენად; Elastic Net სასარგებლოა რეგულარიზაციისა და მნიშვნელოვანი ცვლადების შერჩევისათვის; ხოლო gradient boosting-ის მეთოდები, მათ შორის XGBoost, უზრუნველყოფს პროგნოზის მაღალი სიზუსტით გენერირებას და მოდელის შედეგების შედარებით მკაფიო ინტერპრეტაციას. თუმცა, დროითი დამოკიდებულებების სრულყოფილად

ასახვისათვის საჭიროა ისეთი არქიტექტურების გამოყენება, რომლებიც უკეთ ამუშავებენ მიმდევრობით მონაცემებს.

ამ თვალსაზრისით განსაკუთრებული მნიშვნელობა ენიჭება ღრმა სწავლების მოდელებს, მათ შორის რეკურენტულ ნეირონულ ქსელებს, LSTM არქიტექტურას, attention-მექანიზმებზე დაფუძნებულ მოდელებსა და ტრანსფორმერულ მიდგომებს. გრძელვადიანი და ალბათური პროგნოზირებისათვის შეიძლება გამოყენებულ იქნეს ისეთი მოდელები, როგორცაა DeepAR და Temporal Fusion Transformer, რომლებიც საშუალებას იძლევა პროგნოზში გათვალისწინებულ იქნეს როგორც დროითი დინამიკა, ისე მაკროეკონომიკურ ცვლადებს შორის ფარული დამოკიდებულებები.

მესამე მიმართულებად შეიძლება განვიხილოთ გაძლიერებით სწავლება, რომელიც პროგნოზირების სისტემაში უმეტესად გამოიყენება არა როგორც პირველადი პროგნოზირების მექანიზმი, არამედ როგორც გადაწყვეტილების კორექციისა და ოპტიმიზაციის ფენა. ასეთი მიდგომა განსაკუთრებით მნიშვნელოვანია იმ შემთხვევებში, როდესაც საჭიროა პროგნოზის საფუძველზე პოლიტიკის, რესურსების განაწილების ან მართვის სტრატეგიის ადაპტირება. განმტკიცებითი სწავლების ელემენტების ინტეგრაცია ზრდის სისტემის შესაძლებლობას, არა მხოლოდ შეაფასოს მოსალოდნელი შედეგი, არამედ შესთავაზოს ოპტიმალური რეაგირების სცენარი.

დამატებით, nudging-ის კონცეფციისა და განმტკიცებითი სწავლების კომბინაცია შეიძლება გამოყენებულ იქნეს ხელოვნური ინტელექტის მხარდაჭერით გადაწყვეტილების მიღების პროცესში. აღნიშნული მიდგომა საშუალებას იძლევა პროგნოზირების შედეგები გარდაიქმნას რეკომენდაციებად, რომლებიც მიმართულია ეკონომიკური პოლიტიკის ან ადმინისტრაციული გადაწყვეტილებების უფრო ეფექტიანად ფორმირებისაკენ.

საბოლოოდ, მაკროეკონომიკური პროგნოზირებისათვის ყველაზე გამართლებულია ჰიბრიდული და ანსამბლური არქიტექტურა, რომელიც აერთიანებს ეკონომეტრიკულ მოდელებს, მანქანური სწავლების ალგორითმებსა და ღრმა ნეირონულ ქსელებს. ასეთი კომბინაცია ამცირებს ერთი კონკრეტული მოდელის შეზღუდვებზე დამოკიდებულებას, აუმჯობესებს პროგნოზის მდგრადობას და ზრდის მოდელის უნარს, იმუშაოს როგორც მოკლევადიან, ისე საშუალოვადიან პროგნოზირების ამოცანებში.

2. პრობლემის ფორმულირება

2.1. მაკროეკონომიკის ზოგადი ფორმალური მოდელი

ნებისმიერი სფეროს ზოგადი ფორმალური მაკროეკონომიკური მოდელი შეიძლება წარმოდგენილი იყოს დინამიკური არაწრფივი განტოლებების შემდეგი სისტემით:

$$\mathbf{y}_{t+1} = \mathcal{F}(\mathbf{y}_t, \mathbf{x}_t) \quad (1)$$

$$\mathbf{z}_t = \mathcal{H}(\mathbf{y}_t, \boldsymbol{\xi}_t) \quad (2)$$

სადაც $\mathbf{y}_t \in \mathbb{R}^{n_y}$ – არის ენდოგენური მაკროეკონომიკური ცვლადების ვექტორი (მაგ., გამოშვება, GDP, ინფლაცია და სხვ.) t მომენტში; $\mathbf{x}_t \in \mathbb{R}^{n_x}$ – წარმოადგენს ეგზოგენურ შეყვანის ვექტორს t მომენტში, რომელიც, სავარაუდოდ, დაკავშირებულია \mathbf{y}_t -თან და რომლის მნიშვნელობებიც განისაზღვრება მოდელის გარეთ არსებული ფაქტორებით და შეიძლება გამოიწვიოს შოკები; $\mathcal{F}: \mathbb{R}^{n_x+n_y} \rightarrow \mathbb{R}^m$ არის არაწრფივი იმპლიციტური ოპერატორი; \mathcal{H} არის დაკვირვების ოპერატორი; $\mathbf{z}_t \in \mathbb{R}^m$ – დაკვირვებადი ცვლადები; $\xi_t \in \mathbb{R}^m$ – ხმაურის ვექტორი; ხოლო $t \in [0, t_c]$ – არის განსაზღვრულ დიაპაზონში ცვლადი დროითი ცვლადი (წელი, კვარტალი, თვე, დღე).

რთულ დინამიკურ სისტემებში, როგორცაა მაკროეკონომიკა, შეყვანისა და გამოსვლის ინდიკატორებს შორის კავშირი შეიძლება უფრო ღრმა იყოს; მოცემულ მომენტში გამოსვლის მახასიათებლის სიდიდე შეიძლება დამოკიდებული იყოს არა მხოლოდ სისტემის წინა მდგომარეობაზე, არამედ უფრო ადრინდელ მდგომარეობებზეც. მრავალ შეფასებაში მნიშვნელოვანია ასეთი დამოკიდებულებების ფორმის განსაზღვრა. პროგნოზირების ამოცანებში ასევე მნიშვნელოვანია ჯგუფური დროითი დამოკიდებულებების განსაზღვრა, მაგალითად, კვარტალური ინდიკატორებისგან შედგენილი წლიური მაჩვენებელი როგორ არის დამოკიდებული წინა პერიოდის ინდიკატორებზე. ასეთ მოდელს შეიძლება ეწოდოს მრავალგანზომილებიანი დროითი მწკრივების მრავალსაფეხურიანი პროგნოზირების მოდელი:

$$[\hat{\mathbf{y}}_{t+1}, \hat{\mathbf{y}}_{t+2}, \dots, \hat{\mathbf{y}}_{t+H}]^T = \mathcal{F}([\mathbf{y}_t, \dots, \mathbf{y}_{t-W}]^T, [\mathbf{x}_t, \dots, \mathbf{x}_{t-W}]^T) + \xi_t \quad (3)$$

$\hat{\mathbf{y}}$ არის დაკვირვებადი მიზნობრივი ცვლადი; H და W შესაბამისად წარმოადგენენ პროგნოზირების ჰორიზონტისა და დაკვირვების ფანჯრის განზომილებებს. ამგვარად დასმული პრობლემები ფართოდ გამოიყენება მართვაში, რისკის შეფასებასა და პროგნოზირებაში და კარგად არის შესწავლილი სამეცნიერო ლიტერატურაში [16], [17]. მაკროეკონომიკური ინდიკატორების პროგნოზირებაში ხმაურის დონე შეიძლება ჩაითვალოს მცირე სასარგებლო ინფორმაციის დონესთან შედარებით, ხოლო ფორმულა (3) შეიძლება გადაიწეროს შემდეგნაირად:

$$\hat{\mathbf{Y}}_{t+H} = \mathcal{F}(\ddot{\mathbf{X}}_t) + \xi_t \quad (4)$$

სადაც $\ddot{\mathbf{Y}}_t \in \mathbb{R}^{m \times H}$ არის პროგნოზული მიზნობრივი გამოსვლის მახასიათებელთა მატრიცა, $\ddot{\mathbf{X}}_t = [\mathbf{X}_t, \mathbf{Y}_t] \in \mathbb{R}^{W \times (n_x+n_y)}$ კი წარმოადგენს ფორმულა (3)-ში განსაზღვრულ შეყვანის ცვლადების მატრიცას. შემდგომში ორი ცვლადის კომბინაციას კონკრეტული მიმართულებით ორ წერტილიანი აღნიშვნით წარმოვადგენთ.

ჩვენი ამოცანის მიზანია ფორმულა (4)-ით აღწერილი და ინტერვალზე ($t_c < t \leq t_c + H$) გაზომილი გამოსვლის ვექტორების $\hat{\mathbf{y}}_t$ მნიშვნელობების პროგნოზირება ისტორიული მონაცემების საფუძველზე, რომლებიც მიღებულია ინტერვალში $[0, t_c]$ ცხადია, $W < t_c$.

\mathcal{F} ოპერატორის იმპლიციტური ბუნება პრობლემის ეფექტურად გადაწყვეტას ართულებს. ამ საკითხის გადასაჭრელად ნაშრომში შემოთავაზებულია ღრმა

ნეირონული ქსელის მოდელი, რომელიც ისტორიულ მონაცემებზე დაყრდნობით ახორციელებს F ოპერატორის მრავალსაფეხურიან ადაპტურ აპროქსიმაციას ჰიბრიდული (LSTM-Attention-Nudging) არქიტექტურის ღრმა ნეირონული ქსელებით.

2.2. შემოთავაზებული მეთოდოლოგიური მიდგომა

შემოთავაზებული მოდელი მოიცავს ორ ქვემოდელს: LSTM-Attention და Nudge ქვემოდელებს. თავდაპირველად მოვამზადოთ ამ მოდელების ფორმალური საფუძველი.

ვივარაუდოთ, რომ გვაქვს ისტორიული მონაცემების ნაკრები $\mathbf{D} = \{\tilde{\mathbf{x}}_t, \tilde{\mathbf{y}}_{t+1} \mid t = \overline{0, N}\}$, რომელიც მოიცავს მაკროეკონომიკური ინდიკატორების დაკვირვებებს დროით ინტერვალში $[0, t_N]$ და წარმოდგენილია N-ელემენტიანი დროითი მწკრივის სახით.

იყოს $\tilde{\mathbf{x}}_t = [\mathbf{x}_t, \tilde{\mathbf{y}}_t] \in \mathbb{R}^n$, სადაც $n = n_x + n_y$. აქ \mathbf{x}_t და $\tilde{\mathbf{y}}_t$ შესაბამისად წარმოადგენენ ეგზოგენურ და ენდოგენურ ცვლადების ვექტორებს დროის t მომენტში, ხოლო $\tilde{\mathbf{y}}_{t+1}$ არის t+1 მომენტში მიზნობრივი დაკვირვებადი ცვლადების მრავალგანზომილებიანი ვექტორი. მაკროეკონომიკური მახასიათებლების შემთხვევაში D არის სასწავლო დროითი მწკრივი, ხოლო ამოცანა წარმოადგენს დროითი მწკრივების მრავალგანზომილებიან, მრავალსაფეხურიან ზედამხედველობით მანქანური სწავლების პროგნოზირების ამოცანას, რომელიც შემდეგნაირად იწერება:

მოცემული მრავალგანზომილებიანი სასწავლო ნაკრების D პირობებში განისაზღვრება სრული მანქანური სწავლების პროგნოზირების ოპერატორი \mathcal{M} , რომელიც ფორმულა (4)-ში განსაზღვრულ F ოპერატორს აახლოებს და სისტემის გამოსვლის ცვლადებს შემდეგი ფორმულით პროგნოზირებს:

$$\{\tilde{\mathbf{y}}_{t+1:t+H}\} = \hat{\mathcal{M}}(\{\tilde{\mathbf{x}}_{t-w:t}\} \mid \boldsymbol{\theta}), \quad t \in [0, t_N] \quad (5)$$

სადაც $\{\tilde{\mathbf{y}}_{t+1:t+H}\}$ არის პროგნოზის მრავალსაფეხურიანი დროითი მწკრივი, ხოლო $\boldsymbol{\theta}$ — მოდელის პარამეტრების ვექტორი.

ფორმულა (5)-ში მოცემული \mathcal{M} ოპერატორის ეფექტიანობა შეიძლება შეზღუდული იყოს, განსაკუთრებით რთულ სისტემებში, სადაც კომპონენტებს შორის დამოკიდებულებები არასაკმარისად არის განსაზღვრული. ამის გადასაჭრელად შემოთავაზებულია მრავალფენიანი არქიტექტურა: საბაზისო ფენა იყენებს LSTM-Attention (L-A) მოდელს წინასწარი ამოხსნისთვის, ხოლო მეორე ფენა, Nudge (N), ადაპტურად არეგულირებს ნასწავლ პარამეტრებს Nudging მეთოდის მეშვეობით თითოეული მიზნობრივი ინდიკატორის პროგნოზის სიზუსტის გასაუმჯობესებლად.

სიცხადისთვის შემოვიტანოთ აღნიშვნა, რომლის მიხედვითაც მრავალგანზომილებიანი ოპერატორები და ცვლადები წარმოდგენილია სამწერტილიანი აღნიშვნით, ხოლო ერთგანზომილებიანი — ერთწერტილიანი აღნიშვნით.

\mathcal{M} ოპერატორის ადაპტური, მრავალფენიანი იმპლემენტაცია წარმოვადგინოთ L-A მოდელის ერთი $\hat{\mathcal{M}}^\theta$ ოპერატორით, რომელსაც აქვს ღრმა ნეირონული ქსელის (DNN)

ნეირონის წონები Θ , და k ერთგანზომილებიანი პროგნოზირების nudging (N) მოდელ-
ოპერატორებით \mathcal{M}_i თითოეული მიზნობრივი მახასიათებლისთვის i . პროგნოზი
შეიძლება ჩაიწეროს შემდეგნაირად:

$$\hat{y}_{t+1}^i = \mathcal{M}_i(\{\hat{\mathbf{x}}_{t-w:t}\} | \Theta_i), \quad i = 0, \dots, k \quad (6)$$

სადაც $\hat{y}_{t+1}^i \in \mathbb{R}$ არის i -ე მიზნობრივი მახასიათებლის პროგნოზი, ხოლო სასწავლო
ნაკრების მწკრივი $\hat{\mathbf{x}}_t = [\mathbf{x}_t, y_t] \in \mathbb{R}^n$, ($n = n_x + 1$), $k = m + q$. ცვლადები m და q
შესაბამისად წარმოადგენენ მიზნობრივი მახასიათებლებისა და კვარტალების
რაოდენობებს. \mathcal{M}_i შეიძლება ჩაიწეროს შემდეგი ფორმულით:

$$\mathcal{M}_i = \hat{\mathcal{M}} \circ \varphi_i \circ \mathcal{B}_i \quad (7)$$

სადაც φ_i არის პროექციის ოპერატორი, რომელიც მახასიათებელთა ვექტორიდან
მიზნობრივ მახასიათებელს განსაზღვრავს. \mathcal{B}_i ოპერატორი ასახავს i -ე კონკრეტული
მახასიათებლის მოდელის იმპლემენტაციას DNN-ით, რომელიც წარმოდგენილია L-A-N
არქიტექტურით.

$$\mathcal{B}_i^{\theta_i} = \mathcal{C}_i \circ \mathcal{N}_i^{\theta_i} \quad (8)$$

სადაც \mathcal{C}_i არის კალიბრაციის ოპერატორი, რომელიც \mathcal{M}_i მოდელის წონებს (bias)
ითვლის $\hat{\mathcal{M}}$ მოდელის წონებითა და \mathcal{N}_i ოპერატორით მიღებული მნიშვნელობების
საფუძველზე:

$$\hat{\theta}_i = \hat{\theta}_i + \alpha_i, \quad \alpha_i = \lambda \times \|\delta\| \quad (9)$$

რაც დამოკიდებულია მონაცემთა nudging-ის, რიცხვითი მოდელისა და რიცხვითი
პროგნოზირების მოდელის პირობით რიცხვებზე, სადაც $\delta = \hat{y}_t - \tilde{y}_t$ არის შეცდომა
პროგნოზულ მახასიათებელსა და მის ნამდვილ მნიშვნელობას შორის. ნასწავლი
კორექცია განისაზღვრება როგორც:

$$y_{t+1}^* = \hat{y}_{t+1} + \mathcal{B}_i^{-1}(\hat{y}_t - y_t) \quad (10)$$

თითოეული i -ე მახასიათებლის პროგნოზირებისთვის ახალი \mathcal{B} მოდელების შექმნა
შეიძლება გაგრძელდეს მანამდე, სანამ $\delta < \delta_{tres}$, ან ნაბიჯების რაოდენობა წინასწარ
განისაზღვრება. L-A-N ქსელის პირველი ფენისთვის დანაკარგის ფუნქცია შეიძლება
წარმოდგენილი იყოს შემდეგნაირად:

$$\mathcal{L}(\theta) = \sum_{t=1}^{N-H} \|\hat{\mathcal{M}}_{\theta}(\hat{\mathbf{x}}_t) - \hat{\mathbf{y}}_t\| \quad (11)$$

შესაბამისად, პარამეტრების ოპტიმიზაციის პრობლემა შეიძლება ჩაიწეროს როგორც
 $\hat{\theta}^* = \operatorname{argmin}_{\theta} \{\mathcal{L}(\theta)\}$. აქ $\|\cdot\|$ განისაზღვრება ან L1 ნორმით, ან L2 ნორმით.

$i=0$ -დან $m \times q$ -მდე თითოეული \mathcal{B}_i -სთვის პარამეტრები $\hat{\theta}_i$ გამოითვლება შემდეგი
ფორმულით:

$$\hat{\theta}_i^* = \operatorname{argmin}_{\theta} \|\mathcal{B}_i^{\theta}(\hat{\mathbf{x}}_{t+1}) - \hat{\mathbf{y}}_t\| \quad (12)$$

სადაც საწყისი მნიშვნელობა $\theta_i^0 = \theta^*$. პარამეტრების გამოთვლის სიზუსტის გასაზრდელად ზემოთ განხილული სქემა შეიძლება ორგანიზდეს მრავალსაფეხურიან ადაპტურ პროცესად. $\mathcal{B}_i^{k+1 \rightarrow \theta_i^{k+1}} \rightarrow \mathcal{B}_i^{k \rightarrow \theta_i^k}$, θ_i^0 პროცესი შეიძლება შეიზღუდოს ნაბიჯების რაოდენობით k_{max} ან შეცდომის ზღვრული მნიშვნელობით δ_{thresh}

3. პრობლემის გადაწყვეტა

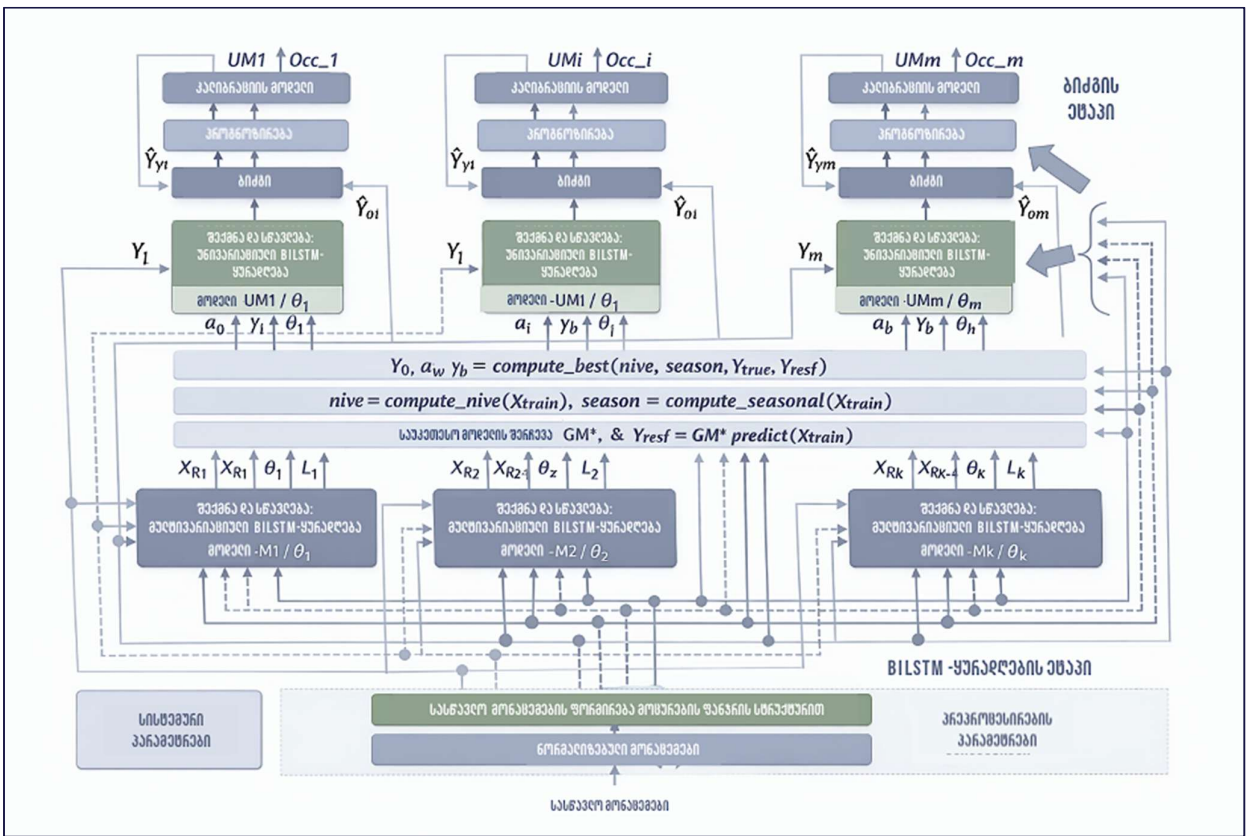
მრავალგანზომილებიანი დროითი მწკრივების მრავალსაფეხურიანი შეფასებისთვის ზემოთ განხილული მეთოდოლოგიის საფუძველზე შემუშავდა ჰიბრიდული მრავალფენიანი, ადაპტური მოდელი, რომლის არქიტექტურა ნაჩვენებია ნახ. 1-ზე. განვიხილოთ თითოეული ფენა ცალ-ცალკე.

3.1. წინასწარი დამუშავების ფენა

პირველი ფენის საწყის ეტაპზე, სხვადასხვა მახასიათებლების მასშტაბების გასათვალისწინებლად და რიცხვითი სტაბილურობის უზრუნველსაყოფად, თითოეული მახასიათებლის ვექტორი უნდა ნორმალიზდეს $[0, 1]$ დიაპაზონში min-max ნორმალიზაციის გამოყენებით. მასშტაბირების პარამეტრები აღებულია მხოლოდ სასწავლო ნაკრებიდან, რათა თავიდან იქნას აცილებული მონაცემთა გაჟონვა.

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (13)$$

ფენის შემდეგ ეტაპზე პროგნოზირების ამოცანა ჩამოყალიბდა როგორც ზედამხედველობითი სწავლების პრობლემა sliding window მიდგომის გამოყენებით, რომელიც მიმდევრობით მონაცემებს ზედამხედველობითი სწავლების ამოცანად გარდაქმნის გადაფარვადი input-output წყვილების შექმნით. დროითი მწკრივების პროგნოზირებაში sliding window მიდგომა მიმდევრობით მონაცემებს ფიქსირებული სიგრძის input-output წყვილებად გარდაქმნის მონაცემთა ნაკრებზე ფანჯრის გადაადგილებით [13]. ეს მეთოდი მოდელს საშუალებას აძლევს ისწავლოს დროითი დამოკიდებულებები წარსული დაკვირვებების (შეყვანის) გამოყენებით მომავალი მნიშვნელობების (გამოსვლის) პროგნოზირებისთვის. ის ფართოდ გამოიყენება რეგრესიაზე დაფუძნებულ მოდელებში და ზრდის პროგნოზირების სიზუსტეს რთული ნიმუშების დაჭერით. ნახ. 2-ზე ნაჩვენებია სასწავლო მონაცემების აგება sliding window-ის გამოყენებით. შედეგად მიიღება ორი მონაცემთა ნაკრები: X_{train} და Y_{train} , სადაც შეყვანისა და გამოსვლის მიმდევრობები აგებულია გადაფარვადი ფანჯრებით.



ნახაზი 1. LSTM-Attention-Nudging მოდელის არქიტექტურა.

3.2. LSTM-Attention ფენა

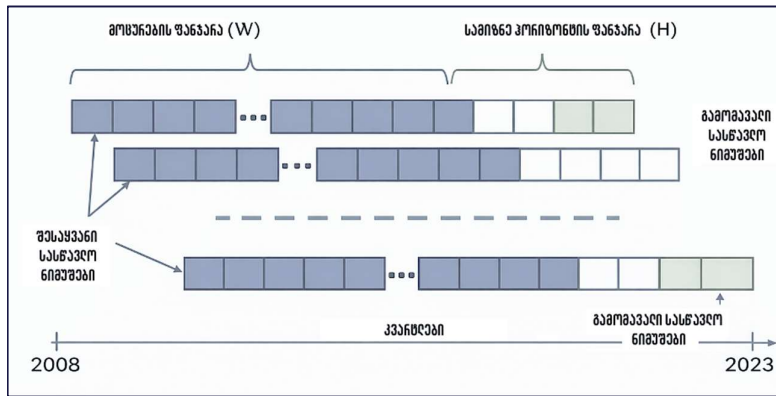
როგორც ნახ. 1-დან ჩანს, ეს ფენა შედგება k პარალელურად მომუშავე LSTM-Attention მოდულისგან. თითოეული მოდული შედგება BiLSTM და Attention მოდულებისგან, რომლებიც სერიულადაა დაკავშირებული. განვიხილოთ ისინი ცალ-ცალკე.

BiLSTM ფენა ამუშავებს შეყვანის მიმდევრობას და ამოიღებს მდიდარ კონტექსტურ მახასიათებლებს როგორც წინამავალი, ისე უკუმიმართული დროითი მიმართულებებიდან. „ორმიმართულებიანობა“ საშუალებას აძლევს ორივე ინტერპრეტაციის ჩაშენებას დამალულ მდგომარეობებში.

BiLSTM და LSTM ქსელების აგება და გამოყენება ფართოდ არის განხილული მრავალ ნაშრომში [18], [19]. ეს ჩვენი კვლევის მთავარი ფოკუსი არ არის. ამიტომ აქ მხოლოდ მოკლედ შევეხებით მათ კავშირებს სხვა ფენებთან.

ეს ფენა წარმოვადგინოთ მიმდევრობების შემდეგი გარდაქმნის სახით:

$$\tilde{\mathbf{h}}_t, \mathbf{h}_t, \mathbf{c}_t = BiLSTM(\mathbf{x}_t, \mathbf{h}_{t-1}, \mathbf{c}_{t-1}) \quad (14)$$



ნახაზი 2. სასწავლო მონაცემების აგების ილუსტრაცია sliding window-ის გამოყენებით: პროგნოზირების ჰორიზონტი $H=4$ კვარტალი და მოძრავი ფანჯრის ზომა $W=k \cdot T_h$.

სადაც h_t არის BiLSTM მოდულის დამალული მდგომარეობა t მომენტში, c_t — უჯრედის მდგომარეობა, ხოლო $\tilde{h}_t = [h_{t-w}, h_{t-2}, \dots, h_{t-1}]$ არის ყველა დამალული მდგომარეობის კონკატენაცია, რომელიც w დროით ნაბიჯზე იქმნება. ისტორიულ მახასიათებლებში რთული დროითი ნიმუშების უკეთ დასაჭერად ვიყენებთ ორ BiLSTM ფენას (რომლებიც აბრუნებენ მიმდევრობებს), additive attention ფენას (რომელიც სწავლობს დროითი მნიშვნელობის წონებს) კონტექსტური ვექტორის მისაღებად და შემდეგ Dense ფენას $horizon \times features$ გამოსვლების წარმოებისთვის.

Attention მოდული განსაზღვრავს, დროის რომელი მომენტები უფრო მნიშვნელოვანია პროგნოზისთვის. იმისათვის, რომ მოდელმა დინამიკურად მოახდინოს ფოკუსირება ყველაზე რელევანტურ დროით ნაბიჯებზე, ინტეგრირებულია Bahdanau-ის ტიპის attention-მექანიზმი [20]. Bahdanau attention-მექანიზმი ითვლის შესაბამისობის ქულებს დეკოდერის წინა დამალულ მდგომარეობასა (query) და ენკოდერის ყველა დამალულ მდგომარეობას (keys/values) შორის. Attention ფენა ითვლის კონტექსტურ ვექტორს ctx , როგორც BiLSTM-ის დამალული მდგომარეობების შეწონილ ჯამს. მექანიზმი ითვლის შესაბამისობის ქულებს e_t , რომლებიც შემდეგ ნორმალიზდება softmax ფუნქციით attention-ის წონების α_t მისაღებად.

$$ctx_t = \sum_{i=t-w}^t \alpha_i h_i \quad (15)$$

ქულის ფუნქცია გამოითვლება LSTM ქსელის პარამეტრებიდან W_h, b_h და აქვს შემდეგი სახე:

$$e_t = v_a^T \tanh(W_a h_t + b_a) \quad (16)$$

სადაც v_a, W_a, b_a — სასწავლო წონითი მატრიცები/ვექტორებია. Attention-ის წონები α_t გამოითვლება feedforward ნეირონული ქსელის გამოყენებით:

$$\alpha_t = \frac{\exp(e_t)}{\sum_{k=1}^T \exp(e_k)} \quad (17)$$

ეს მოდელს საშუალებას აძლევს დინამიკურად ფოკუსირდეს შეყვანის მიმდევრობის რელევანტურ ნაწილებზე და მიიღოს ისტორიის შეწონილი წარმოდგენა.

LSTM-Attention ქსელის შეჯამება გადაეცემა საბაზისო პროგნოზირების ქვეფენას, რომელიც წარმოდგენილია Dense კვანძებით აგებული MLP-ის სახით.

საბოლოო ოპერაციაა პროგნოზის residual და naïve მნიშვნელობების შეფასება. Naive baseline უბრალოდ იმეორებს ბოლო დაკვირვებულ ვექტორს თითოეული მომავალი ნაბიჯისთვის (ძლიერი და მარტივი baseline მრავალი რეალური მწკრივისთვის) და წარმოდგენილია შემდეგი ფორმით:

$$\hat{y}_{t:t+H}^n = \hat{y}_t \quad (18)$$

Residual baseline გამოითვლება როგორც:

$$\hat{y}_{t:t+H}^r = \hat{y}_{t:t+H} - \hat{y}_{t:t+H}^n \quad (19)$$

სადაც $\hat{y}_{t-H:t}$ არის მიზნობრივი მახასიათებლების პროგნოზი, რომელიც აგებული მოდელით შესრულდა. მოდელი end-to-end რეჟიმში სწავლობს საშუალო აბსოლუტური შეცდომის (MAE) დანაკარგის მინიმიზაციით Adam ოპტიმიზატორის გამოყენებით. პროგნოზი გაკეთდა საბოლოო სასწავლო ფანჯრისთვის და შეფასდა residual და naïve მნიშვნელობები.

რადგან ღრმა ნეირონული ქსელების ოპტიმიზაციის შედეგი მნიშვნელოვნად დამოკიდებულია პარამეტრების საწყისი მნიშვნელობების არჩევაზე, ამ ნაშრომში ერთდროულად აირჩევა v რაოდენობის LSTM-Attention ნეირონული ფენა და ოპტიმიზაცია მათზე დამოუკიდებლად, შემთხვევით არჩეული საწყისი მნიშვნელობებით სრულდება. შესაბამისად, ერთი შედეგის ნაცვლად მიიღება v შედეგი.

i -ე BiLSTM-Attention მოდელის განშტოების მილსადენი შემდეგნაირად გამოიყურება:

$$(X, Y)_{train}, \theta_0 \rightarrow LSTM - Attention_i \rightarrow (ctx, h^*)_i \rightarrow MLP_i \rightarrow (\hat{y}, \hat{y}^r, \hat{y}^n, \mathcal{A})_i \quad (20)$$

შემდეგი ნაბიჯი გულისხმობს ყველა v მილსადენიდან ოპტიმალური განშტოების შერჩევას და მის დაყვანას შერეული მოდელების წრფივ ინტერპოლაციამდე. ამისთვის განხორციელდა შემდეგი ნაბიჯები: I) შერჩეული v მოდელიდან ირჩევა მაქსიმალური პროგნოზირების სიზუსტის მქონე მოდელი; II) გამოითვლება მიზნობრივი ცვლადების პროგნოზული residual მნიშვნელობების საშუალო- $\bar{y}_{t-H:t}^r = \frac{1}{v} \sum_{i=1}^v \hat{y}_{i,t-H:t}^r$; III) თითოეული მიზნობრივი f ინდიკატორისა და q -ე სეზონური მახასიათებლისთვის მნიშვნელობა გამოითვლება საშუალო კვადრატული გადახრის მიხედვით; IV) შერჩეული α და γ კოეფიციენტებით residual პროგნოზი მაქსიმალურად უახლოვდება მიზნობრივი მახასიათებლების ნამდვილ მნიშვნელობებს; V) აგებული მოდელების პარამეტრები, წონები და შეფასებები ინახება შემდგომი ფენების გამოთვლებში უკუკავშირისთვის.

LSTM-Attention მოდელის თითოეული განშტოებიდან მიღებული გამოსავალი გადაეცემა nudging ფენას და იქმნება ერთიანი განშტოება. განვიხილოთ nudging ფენის ფუნქციონალი.

3.3. Nudging ფუნა

ერთგანზომილებიანი მრავალსაფეხურიანი პროგნოზისთვის nudging მოდელის კორექციის პროცესი წარმოვადგინოთ შემდეგი Algorithm 1-ის სახით. **Algorithm 1:** Nudging correction for the i-th branch in the univariate multiple-step prediction

1: Input: θ_1^i - LSTM-Attention network weights, X^i, y^i – last window train data, \mathcal{A}_1^i - accuracy prediction from the previous stage, \mathcal{A}^{tres} –threshold of the accuracy for all branches.

2: Outcome: Corrected LSTM-Attention network weights, \mathcal{A}_k^i - accuracy;

3: $k \leftarrow 1$;

4: if $\mathcal{A}_k^i \geq \mathcal{A}^{tres}$: **return** $\theta_k^i, \mathcal{A}_k^i$;

5: Initialize univariate LSTM-Attention model M_k with parameters θ_k^i ;

6: Compile M_k ;

7: $\theta_k^i \leftarrow \text{Train}(M_k, X^i, y^i)$;

8: $X_T^i \leftarrow \text{LastWindow}(X^i)$;

9: $\check{y}_e^i \leftarrow \text{Predict}(M_k, X_e^i)$; # predict for the last window

10: $\check{y}_{e-1}^i \leftarrow \text{GetFromHistory}(X^i, y^i)$;

11: $\hat{y}_n^i \leftarrow \text{CalculateNudgePrm}(\check{y}_e^i, \check{y}_{e-1}^i, y_e^i, \mathcal{A}_k^i)$;

12: $\theta_k^i \leftarrow \text{CalibrateModel}(M_k, \check{y}_e^i, \hat{y}_n^i)$

13: $\mathcal{A}_k^i \leftarrow \text{CalculateAccuracy}(\hat{y}_e^i)$;

14: $k \leftarrow k+1$;

15: goto 4;

აქ ყურადღებას გავამახვილებთ შემოთავაზებული ალგორითმის მხოლოდ ორ პუნქტზე — 11-ე და 13-ე პუნქტებზე.

3.3.1. Nudging პარამეტრების გამოთვლა

განვიხილოთ მოქმედებები თანმიმდევრობით.

1. ნაჯინგის კორექცია გამოითვლება შერჩეული პოლიტიკებისა და სისტემური პარამეტრების შესაბამისად:

$$\check{y}_{\text{nudged}} = \begin{cases} \check{y}_e - \delta, & \mathcal{A}_k \geq \mathcal{A}_{\text{thresh}} \ \& \ y_e < \check{y}_e \\ \check{y}_e + \delta, & \mathcal{A}_k \geq \mathcal{A}_{\text{thresh}} \ \& \ y_e > \check{y}_e \end{cases}$$

სადაც $\delta = \text{FEEDBACK_SHIFT}$ არის სისტემური პარამეტრი.

2. სეზონურობის კომპონენტი y_{season} გამოითვლება ისტორიული მონაცემების საფუძველზე.
3. მიზნობრივი მნიშვნელობა გამოითვლება საბაზისო და სეზონური კომპონენტების შეწონვით:

$$\hat{y}_{\text{nudged}} = (1 - \lambda) \times \hat{y}_{\text{nudged}} + \lambda \times y_{\text{season}}, \quad \lambda \in [0,1].$$

4. პროგნოზის ნდობის რეგიონი $\text{tr} = [c_{\min}, c_{\max}]$ ყალიბდება წინა პროგნოზის შესაბამისად.
5. მიზნობრივი მნიშვნელობა იზღუდება ნდობის ინტერვალის შესაბამისად:

$$\hat{y}_{\text{nudged}} = \text{clip}(\hat{y}_{\text{nudged}}, \text{tr}).$$

3.3.2. მოდელის კალიბრაცია

კალიბრაცია ხორციელდება შემდეგი ნაბიჯებით:

2. პროგნოზირებულ და ნაჯინგით კორექტირებულ მახასიათებლებს შორის გადახრა გამოითვლება შემდეგნაირად:

$$\Delta = \hat{y}_e^i - \hat{y}_h^i.$$

4. მოდელიდან აღებულია წონები და bias პარამეტრები.
5. ხორციელდება bias პარამეტრის კორექტირება:

$$\text{bias} = \text{bias} + \Delta.$$

ალგორითმი 1 უზრუნველყოფს პროგნოზირების პროცესის ადაპტირებადობას, რადგან თითოეული მახასიათებლის პროგნოზირებისას k -ნაბიჯების რაოდენობა შეიძლება განსხვავებული იყოს და იგი შეიძლება დარეგულირდეს როგორც ავტომატურად, ისე ხელით.

4. ექსპერიმენტი

4.1. მონაცემთა ნაკრებები

პრაქტიკული პროგნოზირების შედეგიანობის შესაფასებლად შემოთავაზებული მოდელი შემოწმდა როგორც სინთეზურ მონაცემებზე, ისე შემდეგ რეალურ მონაცემთა ნაკრებებზე: 1) საქართველოს კვარტალური მაკროეკონომიკური მონაცემთა ნაკრები (2008–2025) geostat.ge-დან, რომელიც შეიცავს 18 ცვლადს, რომელთაგან რვა იყო პროგნოზირების მიზნობრივი ცვლადი;

4.2. ჰიპერპარამეტრები

მოდელისთვის ჰიპერპარამეტრების არჩევა მნიშვნელოვნად მოქმედებს მის ეფექტიანობასა და სიზუსტეზე. ჰიპერპარამეტრების ოპტიმალური შერჩევითვის გამოვიყენეთ Keras tuner, რომელიც TensorFlow-სთან ერთად სთავაზობს ჩაშენებულ ძიების ალგორითმებსა და ინსტრუმენტებს განსაზღვრულ სივრცეში საუკეთესო

ჰიპერპარამეტრების ეფექტურად მოსაძებნად, რითაც აუმჯობესებს მოდელის სიზუსტესა და განზოგადებას.

პარამეტრი	მნიშვნელობა
მიმდევრობის სიგრძე (L)	128
ბიჯი (S)	1
დამალული განზომილება (თითო BiLSTM ფენაზე)	64
BiLSTM ფენების რაოდენობა	2
Dropout-ის ალბათობა	0.2
ოპტიმიზატორი	Adam
სწავლის სიჩქარე	0.001
Batch Size	64
ეპოქების რაოდენობა	50
Sliding windows პროგნოზის ჰორიზონტის ზომა	4
პროგნოზის ჰორიზონტის ზომა	4

ცხრილი 1. ჩვენი ჰიბრიდული მოდელის სწავლებისთვის გამოყენებული ჰიპერპარამეტრების მნიშვნელობები.

ცხრილი 1 ასახავს საქართველოს ბიუჯეტის პროგნოზირების მაგალითისთვის გამოყენებულ ჰიპერპარამეტრებს. Nudging ფენისთვის FEEDBACK_SHIFT სისტემური პარამეტრი არჩეულია 0.15. ჰიპერპარამეტრების უმეტესობა მონაცემთა ნაკრებებს შორის უცვლელი დარჩა, გარდა ეპოქების რაოდენობისა.

4.3. ექსპერიმენტული დიზაინი

პირველი ექსპერიმენტი იყენებს საქართველოს კვარტალურ მაკროეკონომიკურ ინდიკატორებს 2008Q1-2023Q4 პერიოდისთვის და პროგნოზირებს 2024Q1-2024Q4 პერიოდს. მიზნობრივი ცვლადები მოიცავს GDP-ს, GDP per capita-ს, imports-ს, state revenues-ს, broad money M3-ს, საშუალო გაცვლით კურსს, დასაქმებულთა რაოდენობას და agricultural production output-ს. შემოთავაზებული მოდელი სამჯერ გაეშვა განმეორებით გაშვებებს შორის სტაბილურობის შესაფასებლად. ამ ექსპერიმენტის დროს მოდელის მონაცემებში არ იყო შეყვანილი 2024 წლის მაკროეკონომიკური მონაცემების მნიშვნელობები და შედარება ხორციელდებოდა სისტემის გრეთ.

მეორე ექსპერიმენტი იყენებს საქართველოს კვარტალურ მაკროეკონომიკურ ინდიკატორებს 2008Q1-2024Q4 პერიოდისთვის და პროგნოზირებს 2025Q1-2025Q4 პერიოდს. მიზნობრივი ცვლადები მოიცავს GDP-ს, GDP per capita-ს, exports-ს, imports-ს, state revenues-ს, broad money M3-ს, საშუალო გაცვლით კურსს. ამ ექსპერიმენტის დროს მოდელის მონაცემებში არ იყო შეყვანილი 2025 წლის მაკროეკონომიკური მონაცემების მნიშვნელობები და შედარება ხორციელდებოდა სისტემის გარეთ.

4.4. შედეგები

შემოთავაზებულმა მოდელმა, 2024 წლის მონაცემების პროგნოზირებისას, სამი თანმიმდევრული გაშვებისას სტაბილური შედეგები აჩვენა: 94.09%, 93.91% და 94.16%. იმავე პირობებში 2025 წლის მონაცემების პროგნოზირებისას, აჩვენა: საშუალო პროგნოზირების სიზუსტე - 92.16%.

მთავარი დასკვნა ის არის, რომ შემოთავაზებული მოდელი კონკურენტუნარიანია და აჩვენებს პროგნოზირების მაღალ სიზუსტეს სხვადასხვა ტიპის და ხმაურის მონაცემების პირობებში. შემოთავაზებულ მოდელს აქვს მკაფიო მონაცემთა მილსადენი, განტოლებებით განსაზღვრული პროგნოზირების ოპერატორი, ფიქსირებული ჰიპერპარამეტრები, განსაზღვრული დანაკარგის ფუნქცია და შეზღუდული კორექციის მექანიზმი. ეს აუმჯობესებს რეპროდუცირებადობას და შედეგებს უფრო შესაფერის ხდის აკადემიური შეფასებისთვის.

ინდიკატორი	გაშვება 1	გაშვება 2	გაშვება 3
GDP (M \$)	94.83	94.28	94.69
GDP PER CAPITA \$	94.49	94.40	94.60
Imports (M \$)	92.57	92.16	92.88
State Revenues (M \$)	92.06	92.36	92.10
Broad Money M3 (M \$)	92.00	92.02	92.00
AV. Exchange USD-\$	97.50	97.11	98.07
Number Of Employees	97.15	96.31	97.29
Agricult. Products Outp. (M \$)	92.10	92.66	91.65
Overall average	94.09	93.91	94.16

ცხრილი 2. საქართველოს 2024 წლის მაკროეკონომიკური მაჩვენებლების პროგნოზირების საშუალო სიზუსტე %.

ინდიკატორი	წელი	კვარტალი	მოდულის პროგნოზი	რეალური მონაცემი	პროგნოზის სიზუსტე %
GDP (MILLION \$)	2025	I	8034.73	7896.90	98.28
GDP (MILLION \$)	2025	II	8841.25	9328.60	94.78
GDP (MILLION \$)	2025	III	9530.19	10334.60	92.22
GDP (MILLION \$)	2025	IV	9601.90	10644.80	90.20
GDP PER CAPITA \$	2025	I	2170.25	2131.70	98.22
GDP PER CAPITA \$	2025	II	2354.25	2492.70	94.45
GDP PER CAPITA \$	2025	III	2590.74	2818.90	91.91
GDP PER CAPITA \$	2025	IV	2594.72	2888.90	89.82
EXPORTS (MILLION \$)	2025	I	1486.14	1397.40	94.03
EXPORTS (MILLION \$)	2025	II	1662.64	1837.10	90.50
EXPORTS (MILLION \$)	2025	III	1940.80	1916.60	98.75
EXPORTS (MILLION \$)	2025	IV	1849.20	2140.00	86.41
IMPORTS (MILLION \$)	2025	I	3954.17	4457.80	88.70
IMPORTS (MILLION \$)	2025	II	4457.53	4522.30	98.57
IMPORTS (MILLION \$)	2025	III	4652.39	4670.40	99.61
IMPORTS (MILLION \$)	2025	IV	4915.95	4961.50	99.08
STATE REVENUES (MILLION \$)	2025	I	2149.70	2363.80	90.94
STATE REVENUES (MILLION \$)	2025	II	2036.13	2439.70	83.46
STATE REVENUES (MILLION \$)	2025	III	2044.40	2640.40	77.43
STATE REVENUES (MILLION \$)	2025	IV	2111.71	2701.10	78.18
BROAD MONEY M3 (MILLION \$)	2025	I	17312.24	17379.30	99.61
BROAD MONEY M3 (MILLION \$)	2025	II	17490.65	18382.10	95.15
BROAD MONEY M3 (MILLION \$)	2025	III	18659.64	20018.50	93.21
BROAD MONEY M3 (MILLION \$)	2025	IV	19006.09	21311.20	89.18

AVERAGE EXCHANGE RATE USD-\$	2025	I	2.76	2.81	92.20
AVERAGE EXCHANGE RATE USD-\$	2025	II	2.85	2.74	91.94
AVERAGE EXCHANGE RATE USD-\$	2025	III	2.77	2.71	91.82
AVERAGE EXCHANGE RATE USD-\$	2025	IV	2.86	2.71	91.81
Overall average					92.16

ცხრილი 3. საქართველოს 2025 წლის მაკროეკონომიკური მაჩვენებლების პროგნოზირების საშუალო სიზუსტე %. კვარტლების მიხედვით

6. დასკვნა და მომავალი სამუშაო

ჩატარებული ანალიზის ფარგლებში შეფასდა ნეირონული ქსელების სხვადასხვა არქიტექტურის ფუნქციური შესაძლებლობები და შეზღუდვები, განსაკუთრებით მაკროეკონომიკური ინდიკატორების პროგნოზირების ამოცანებთან მიმართებით. მიღებული შედეგების საფუძველზე გამოიკვეთა, რომ **LSTM-Attention-Nudge** ტიპის ჰიბრიდული მიდგომა აღნიშნულ კონტექსტში მაღალი პრაქტიკული ეფექტიანობით ხასიათდება. ამ არქიტექტურულ საფუძველზე შესაძლებელია პროგნოზირების ახალი სისტემის ფორმირება, რომელიც ერთდროულად ითვალისწინებს დროით დინამიკას, ყურადღების მექანიზმებზე დაფუძნებულ წონადაზიანებას და კორექციული ზემოქმედების კომპონენტს.

სინთეზურ და რეალურ მონაცემებზე განხორციელებულმა ექსპერიმენტულმა ტესტირებამ აჩვენა მოდელის მაღალი პროგნოზული სიზუსტე და პარამეტრული ცვლილებების მიმართ შედარებითი მდგრადობა. ამასთან, გამოვლინდა ის მიმართულებები, რომელთა შემდგომი განვითარებაც შესაძლებელს გახდის მოდელის განმარტებადობის, სტაბილურობისა და გამოყენებითი ღირებულების დამატებით გაძლიერებას.

შემდგომი კვლევითი და ტექნიკური ნაბიჯები:

- მოდელის მათემატიკური საფუძვლების უფრო ღრმა ფორმალიზება და თეორიული ვალიდაციის გაძლიერება.
- შესაბამისი საკორექციო სიგნალების გამოყენება **Nudge** კომპონენტის პარამეტრების ოპტიმალური კალიბრაციისათვის.

• პროგნოზირების არქიტექტურაში დარგობრივად სპეციფიკური, სექტორული და სტრუქტურული შეზღუდვების ინტეგრირება მოდელის ახსნადობისა და სანდოობის გასაზრდელად.

ლიტერატურა

1. Ammara Laeeq, Jie Li, Usman Adeel, A Hybrid LSTM–Attention Model for Multivariate Time Series Imputation: Evaluation on Environmental Datasets, *Mach. Learn. Knowl. Extr.* 2026, 8, 18, <https://doi.org/10.3390/make8010018>
2. Dominik Stempien, and Robert Slepaczuk, Hybrid Models for Financial Forecasting: Combining Econometric, Machine Learning, and Deep Learning Models, *arXiv:2505.19617v1 [q-fin.TR]* 26 May 2025.
3. Kehang Feng, Yuxin Zhang, LSTM-Attention: A Novel Deep Learning Model for Sustainable Development Goals Achievement Prediction, *MLSC '25: Proceedings of the 2025 International Symposium on Machine Learning and Social Computing*, Pages 46–50, <https://doi.org/10.1145/3778450.3778457>
4. Navid Hajizadeh, S. Yazdani, Sara Ershadi-Nasab, Attentional Bi-LSTM for Multivariate Time Series Forecasting on Edge Devices: A Case Study on NanoPi Neo Plus2, DOI:10.1109/ICCKE68588.2025.11273888
5. Inna Polichtchouk, Simon Lang, Sarah-Jane Lock, Michael Maier-Gerber, Peter Dueben, Hybrid ensemble forecasting combining physics-based and machine-learning predictions through spectral nudging, 5 Mar 2026, <https://doi.org/10.48550/arXiv.2603.05570>
6. Jiangcheng Zhu, Shuang Hu, Rossella Arcucci, Chao Xu, Jihong Zhu, and Yi-ke Guo, Model Error Correction in Data Assimilation by Integrating Neural Networks, *BIG DATA MINING AND ANALYTICS*, Volume 2, Number 2, June 2019, DOI: 10.26599/BDMA.2018.9020033
7. Heaney, C. E., Tang, J., Yan, J., Guo, D., Ipock, J., Kaluvakollu, S., et al. (2024). Data assimilation with machine learning for dynamical systems: Modelling indoor ventilation. *Physica A*, 643, 129783. <https://doi.org/10.1016/j.physa.2024.129783>
8. Sun, H., Lei, L., Liu, Z., Ning, L., & Tan, Z. M. (2025). An online paleoclimate data assimilation with a deep learning-based network. *Journal of Advances in Modeling Earth Systems*, 17(6).
9. Jaemin Oh, Jinsil Lee, Youngjoon Hong, Machine Learning-Based Nonlinear Nudging for Chaotic Dynamical Systems, 7 Aug 2025, <https://doi.org/10.48550/arXiv.2508.05778>
10. Harbir Antil, Rainald Löhner, Randy Price, Data Assimilation with Deep Neural Nets Informed by Nudging, 22 Nov 2021, <https://doi.org/10.48550/arXiv.2111.11505>
11. Yingjun Peng, Yulong Bai, Xufeng Wang, Qinghe Yu, Xiaoxin Yue, Nudging-Based Data Assimilation Method for Error Correction Coupled With Huber Loss Functions and BiLSTM-GRU Hybrids, *Journal of Advances in Modeling Earth Systems*, DOI: 10.1029/2025MS005306

12. Hassanzadeh, P., Chattopadhyay, A., Bach, E., Nabizadeh, E. (2023): Deep learning-enhanced ensemble-based data assimilation for high-dimensional nonlinear dynamical systems, <https://doi.org/10.57757/IUGG23-3118>
13. Defu Cao, Yixiang Zheng, Parisa Hassanzadeh, Simran Lamba, Xiaomo Liu, and Yan Liu. Large Scale Financial Time Series Forecasting with Multi-faceted Model. ICAIF '23, <https://doi.org/10.1145/3604237.3626868>
14. Ta-Chung Chi, Ting-Han Fan, Raffaele M. Ghigliazza, Domenico Giannone, Zixuan (Kevin) Wang, Macroeconomic Forecasting and Machine Learning, arXiv:2510.11008v1 [econ.EM] 13 Oct 2025.
15. Cook, Thomas R., and Aaron Smalter Hall. “Macroeconomic Indicator Forecasting with Deep Neural Networks.” Federal Reserve Bank of Kansas City, Research Working Paper 17-11, <https://doi.org/10.18651/RWP2017-11>
16. Qi Tang, Ruchen Shi, Tongmei Fan, Yidan Ma, and Jingyan Huang, Prediction of Financial Time Series Based on LSTM Using Wavelet Transform and Singular Spectrum Analysis, Mathematical Problems in Engineering, 2021, <https://doi.org/10.1155/2021/9942410>
17. Bryan Lima, Sercan O. Arik, Nicolas Loeff, Tomas Pfister, Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting, arXiv:1912.09363v3 [stat.ML] 27 Sep 2020.
18. Zurab Bosikashvili, Tamar Bosikashvili, Ketevan Bosikashvili - Using Blocking Meta-heuristics in the Models of Deep Learning, International scientific practical conference modern challenges and achievements in information and communication technologies – 2023, ISBN 978-9941-512-06-, <http://www.gtu.ge>.
19. Pierre Beaujuge, Multivariate Time Series Forecasting with a Bidirectional LSTM: Building a Model Geared to Multiple Input Series, Medium, Sep 24, 2020.
20. Gholam Reza Alipour Modab, Attention-driven deep learning models for multivariate time series forecasting of reservoir water levels, Water Sci Technol (2025) 92(12): 1650–1668. <https://doi.org/10.2166/wst.2025.182>