

МОДЕЛ ВРЕМЕНСКЕ СЕРИЈЕ ЗА ПРЕДВИЂАЊЕ ЦЕНА AWS СПОТ ИНСТАНЦИ TIME SERIES MODEL FOR PREDICTING AWS SPOT INSTANCE PRICES

Милица Прањкић, Јелена Сливка, *Факултет техничких наука, Нови Сад*

Област – РАЧУНАРСТВО И АУТОМАТИКА

Кратак садржај – *Растућа потражња за рачунарским ресурсима подстакла је коришћење провајдера рачунарства у облаку. То је довело до развоја нове димензије у којој везе између употребе ресурса и трошкова морају бити разматране са организацијске перспективе. Као део своје EC2 (Elastic Compute Cloud) услуге, Амазон је увео спот инстанце као јефтину јавну инфраструктуру која долази са ценом непоузданости услуге. Цена спот инстанци је кључни фактор у управљању трошковима и оптимизацији ресурса у рачунарству у облаку. Динамичко одређивање цена може закомпликовати процес планирања и алокације ресурса. Стога, тачна предикција цена спот инстанци може бити од велике користи за кориснике облака, омогућавајући им да максимално искористе своје ресурсе и минимизирају трошкове.*

Кључне речи: *спот инстанце, временске серије, модел за предвиђање цене, SARIMA.*

Abstract – *The growing demand for computing resources has spurred the use of cloud computing providers. This has led to the development of a new dimension where the relationship between resource usage and cost must be considered from an organizational perspective. As part of its EC2 (Elastic Compute Cloud) service, Amazon introduced spot instances as a cheap public infrastructure at the price of service unreliability. The price of spot instances is a key factor in cost management and resource optimization in cloud computing. Dynamic pricing can complicate the process of planning and resource allocation. Therefore, accurate prediction of spot instance prices can greatly benefit cloud users, allowing them to maximize their resources and minimize costs.*

Keywords: *spot instances, time series, price prediction model, SARIMA*

1. УВОД

У Амазоновом моделу спот инстанци (*Spot Instance, SI*), изнајмљене инстанце могу се нагло прекинути од стране провајдера услуга када је то потребно. Интерфејс за управљање *SI* базира се на стратегији лицитације која се базира на Амазоновим тајним ценовним стратегијама, што отежава корисницима примену стратегија распоређивања или пружања ресурса заснованих на таквим (јефтинијим) ресурсима.

НАПОМЕНА:

Овај рад проистекао је из мастер рада чији ментор је била др Јелена Сливка, ванр. проф.

Иако се верује да употреба *EC2 SI* инфраструктуре може смањити трошкове за крајње кориснике, детаљни преглед литературе показује да њихове карактеристике и могућности још увек нису дубоко истражене.

AWS пружа могућност прегледа историјских података о промени цена спот инстанци у претходних 90 дана. Ова историја, корисницима пружа детаљан увид у флукуацију цена спот инстанци током одређеног периода. У оквиру овог истраживања разматран је процес формирања одговарајућег модела за одређивање цена *SI* на основу прикупљених историјских података. Циљ модела је да омогући предвиђање будуће цене *SI* током времена, како би се лоцирао *SI* ресурс доступан по ниској цени.

Овај рад презентује *SARIMA (Seasonal Autoregressive Integrated Moving Average)* модел временске серије за предикцију цена *SI* у будућности. Скуп података над којим је модел обучен су историјски подаци о ценама *SI* у претходних шест месеци.

Историјски подаци су подељени у тренинг и тест скуп, где се тренинг подаци користе за обуку модела, а тест подаци за његову евалуацију. За процену тачности предвиђања коришћене су различите метрике које пореде стварне вредности цене *SI* у датом временском тренутку са њеном вредношћу предвиђеном од стране модела временске серије.

Остатак овог рада организован је на следећи начин. Поглавље 2 даје преглед радова који се баве предикцијом цена *SI*. Поглавље 3 излаже теоријске основе примењених модела временских серија. Поглавље 4 описује решење за предикцију цена *SI* представљено у овом раду. Поглавље 5 описује експерименте извршене у циљу евалуације овог решења, а резултати евалуације су приказани и дискутовани у поглављу 6. Поглавље 7 закључује рад.

2. ПРЕГЛЕД СТАЊА У ОБЛАСТИ

Од објављивања спот инстанце на Амазону 2009. године, многе студије су се бавиле њиховим истраживањем. Лиу и други [1] су ове студије поделили у две главне категорије, на основу метода на којима се базирају:

- модели за предвиђање цена *SI* на основу статистичке временске анализе (поглавље 2.1.1) и
- модели за предвиђање цена *SI* применом машинског учења (поглавље 2.1.2).

2.1 Модели предвиђања цена на основу статистичке временске анализе

У раду [2], *SARIMA* модел се успоставља анализирајући историјске цене спот инстанце за претходних 11 месеци. Применом различитих хеуристика за моделовање тренинг скупа података, откривено је да *SARIMA* алгоритам постиже значајно боље перформансе приликом предвиђања цене *SI* у односу на размотрене алтернативе. *SARIMA* модел временске серије показује 17% већу тачност у поређењу са резултатима других предиктивних алгоритама.

2.2 Модели предвиђања цена на основу машинског учења

У једном од скоријих истраживања, Конг и други [3] износе став да методе машинског учења нису добре за предвиђање циљне варијабле за проблеме временских серија. *RNN* (*Recurrent Neural Network*) постиже добре перформансе на проблемима временских серија, али тренирање овог модела проблематично због појаве проблема експлозије градијента и нестанка градијента. *LSTM* (*Long Short-Term Memory network*) је позната варијанта *RNN*-а која решава проблем нестанка градијента, експлозије градијента и недостатка дугорочне способности меморије.

Стога се *LSTM* може ефективно користити за временске серије. *GRU* (*Gated Recurrent Unit*) је варијанта *LSTM*-а која поједностављује унутрашњу *LSTM* структуру, а постиже готово једнаке перформансе.

У раду [3], додато је неколико потпуно повезаних слојева у циљу побољшања тачности *GRU* модела у предвиђању нелинеарних података о ценама. Такође, коришћен је модел насумичне шуме (*Random Forest*), метода машинског учења, за предвиђање података о ценама *SI*, ради поређења.

2.2 Одабир модела за предвиђање цене спот инстанце

Преглед стања у области указује на значајне предности алгоритама машинског учења за предикцију временских серија. Међутим, статистички модели временских серија, попут *SARIMA* модела, имају боље перформансе у случајевима [6]:

- Када су подаци имају ограничен опсег вредности, при чему су те вредности ниске
- Када је количина података за обуку мала
- За краткорочна предвиђања
- У случају линеарних временских серија.

Због велике заступљености статистичких алгоритама машинског учења у предикцији цена *SI* у претходним радовима, овај рад ће се фокусирати на истраживање перформанси статичких модела. Такође, количина историјских података над којом ће се вршити обука није велика, правиће се краткорочна предвиђања и цене *SI* су релативно мале вредности, ограничене одређеним опсегом.

Стога, статички модели машинског учења представљају добар избор за овај проблем. Једини услов који историјски подаци не задовољавају, у овом случају, јесте линеарност. Историјски подаци о

ценама инстанци су нелинеарна структура, где сваки податак не може бити објашњен на основу претходних података.

Добијени резултати биће евалуирани на сличан начин као у радовима [3] и [5]. Такође, резултати овог истраживања ће бити упоређени са резултатима из наведених истраживања.

3. ТЕОРИЈСКИ ПОЈМОВИ И ДЕФИНИЦИЈЕ

Како се цене спот инстанци мењају током одређеног временског периода, класификоване су као временска серија (*time-series*). Временска серија је секвенцијални скуп података, мерених у узастопним временским периодима. Мерења прикупљена током догађаја у временској серији сортирају се у одговарајући хронолошки редослед.

Један од најпопуларнијих и најчешће коришћених стохастичких модела временских серија је модел *ARIMA* (*Autoregressive Integrated Moving Average*). За прогнозирање сезонских временских серија, *Бокс* и *Џенкинс* [4] су предложили прилично успешну варијанту *ARIMA* модела, *SARIMA* (*Seasonal ARIMA*).

Бокс-Џенкинс приступ изградњи модела временских серија састоји се од три фазе, које се понављају док се не добије задовољавајући модел: спецификација модела, процена параметара и дијагностичка провера.

У овом раду коришћен је *SARIMAX* модел. *SARIMAX* се разликује у односу на *SARIMA* модел по могућности укључивања утицаја спољних фактора на посматрани модел. Спољне или егзогене варијабле су оне на које се не може директно утицати, али могу имати ефекат на понашање посматраног модела.

4. МЕТОДОЛОГИЈА

Ово поглавље представља имплементацију модела временске серије за предикцију цена *AWS SI*. За формирање и обучавање модела коришћени су историјски подаци о цени *SI* у периоду од 6 месеци. Очекивани излаз из модела је будућа цена *SI*. Као модел машинског учења се користи *SARIMAX* модел.

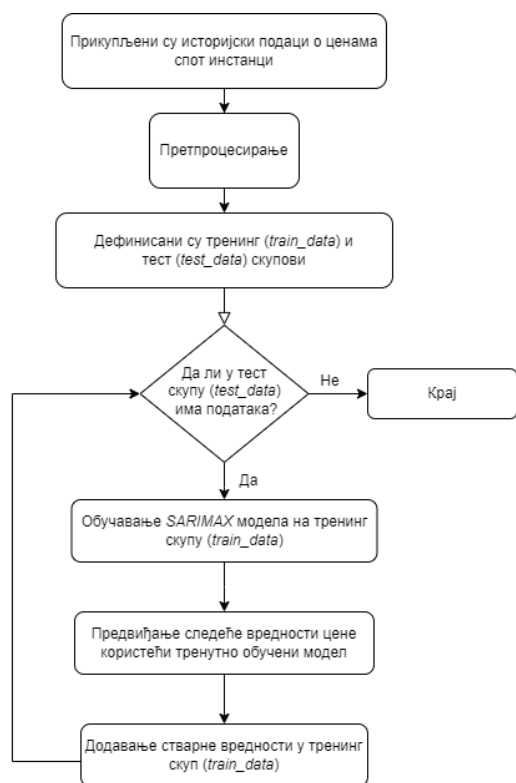
У првом кораку, формира се скуп од почетних тренинг података, који обухвата првих неколико тачака временске серије и користи се за обуку модела.

Затим, тренирани модел предвиђа наредну вредност из временске серије, тј., вредност која следи одмах након последње вредности у подацима за обуку.

Након тога, вредност коју је модел покушао да предвиди, додаје се у скуп тренинг података – не додаје се предвиђена, већ стварна, измерена вредност.

Скуп тренинг података је тиме проширен да обухвата једну додатну тачку.

Овај поступак се понавља за сваку преосталу тачку у временској серији, чиме се добија итеративни поступак обуке који побољшава своје перформансе у сваком кораку (слика 1).



Слика 1. Дијаграм тока експеримента

5. ЕКСПЕРИМЕНТИ

5.1 Скупови података

Историјски подаци о ценама спот инстанци преузети су са *AWS* веб странице. Историја цена доступна је за период од 90 дана. Подаци су прикупљени у два наврата, у размаку од 3 месеца, те скуп података обухвата историју цена у периоду од 180 дана. Посматрани регион је: *us-east-1*, укључујући све његове подрегионе и све типове *SI* доступне у том региону.

Структура прикупљених података представљена је следећим обележјима: зона доступности, тип инстанце, платформа инстанце, цена инстанце и временска ознака узорковања. С обзиром на то да цена уме знатно да варира у зависности од сва три параметра, извршено је груписање скупа података према следећим параметрима: *us-east-1a* (зона доступности), *t3.2xlarge* (тип инстанце), и *Windows* и *Linux* (платформа). Као резултат овог груписања формирана су два засебна скупа података који садрже информације о промени цене спот инстанце која задовољава одабране параметре груписања.

Подаци из поменутих скупова су сортирани у растућем редоследу на основу временске ознаке узорковања, што значи да су историјске цене инстанци поређане хронолошки. Затим, подаци су груписани у тренинг (90%) и тест скуп (10%).

5.2 Експеримент 1

У оквиру овог експеримента, *SARIMAX* моделу допремљен је цео скуп тренинг података над којима је вршена обука. Примарни циљ овог експеримента је, било креирање модела предикције цена *SI*.

Секундарна ствар која се кроз овај експеримент може показати јесте начин на који се развија тачност предвиђања модела током времена. Другим речима, утврдити да ли се квалитет предвиђања побољшава или погоршава како време одмиче.

Након што је модел обучен, извршено је предвиђање цена за неколико тачака у будућности, тј., неколико наредних корака. Број корака је одабран насумично, али пракса је показала да, што се више тачака одабере, то ће се боље приказати и објаснити реално стање система и квалитета дугорочне предикције.

5.3 Експеримент 2

У оквиру овог експеримента, коришћена је техника померајућег прозора за предвиђање. Другим речима, предвиђање је вршено један корак унапред, након чега је резултат тог предвиђања увршћен у тренинг скуп, а модел је поново подвргнут поступку обуке. Овај процес је мање захтеван за меморију, али, обучавање модела траје знатно дуже. Изузетно је ефикасан за предвиђање временских серија и састоји се из неколико корака описаних у поглављу 4.

Након што је поступак обуке модела завршен, добијена су два скупа: скуп предвиђених вредности и скуп стварних вредности. Како би се установила ефикасност модела, ова два скупа су упоређена и над њима су примењене одређене метрике, наведене у наставку (поглавље 5.3).

6. РЕЗУЛТАТИ И ДИСКУСИЈА

У оквиру првог експеримента, модел је обучен над читавим тренинг скупом, а након завршетка процеса учења вршена је предикција будућних вредности. Предикција је вршена за 20 наредних корака.

На почетку, модел показује релативно добре способности предикције. Међутим, како се временски интервал предикције повећава, тако модел све више губи на тачности и показује све лошије перформансе.

Вредности средње апсолутне процентуалне грешке (*MAPE*), приказане у табели 1, представљају просечну грешку модела при предвиђању у односу на стварне вредности. Грешка је изражена у процентима, а почетна вредност код првог корака је 0.0869%, што означава релативно добру прецизност предвиђања у том тренутку. Међутим, како време одмиче и како се повећава број корака у предвиђању, модел почиње да губи своју прецизност. На пример, у 20. кораку процентуална грешка је достигла 2.6332%.

Табела 1. Вредности *MAPE* за различите кораке предикције у експерименту 1

Корак	<i>MAPE</i>
1.	0.0869
2.	0.0324
...	...
19.	2.2682
20.	2.6332

Експеримент 2с, који користи технику померајућег прозора описану у поглављу 4, извршава предикцију

за све вредности из тест скупа. На почетку, модел показује врло добре способности предикције (Табела 2). Тачност модела варира током овог поступка, али апсолутна вредност грешке остаје врло мала до самог краја. Шта више, грешка се смањује како време одмиче.

У експерименту 2, грешка коју модел прави врло је мала од самог почетка. Иако постоје одређене варијације *MAPE* вредности кроз итерације, модел је био у стању да одржи добру тачност у великом броју случајева. Поред тога, тачност овог модела се повећава како време одмиче, што може деловати обећавајуће када је реч о дугорочним предикцијама.

Табела 2. Вредности *MAPE* за различите кораке предикције у експерименту 2с

Корак	<i>MAPE</i>
1.	0.0869
2.	0.1195
...	...
61.	0.0174
62.	0.0398

6.1 Дискусија

Из Табеле 3 јасно се уочава да је техника померајућег прозора правила мање грешке у односу на модел обучен над тренинг који даје предикције за више временских корака у напред.

Табела 3. Вредности *MSE*, *RMSE*, *MAE* и *R-SQUARED* за дату модел

	<i>MSE</i>	<i>RMSE</i>	<i>MAE</i>	<i>R</i> ²
<i>SARIMAX</i>	1.2e-05	0.0035	0.00261	-14.4864
Техника померајућег прозора	2.6e-07	0.0005	0.00042	0.98588

Резултати овог истраживања потврђују теорију изложу у раду [5], која истиче губитак перформанси предикције *SARIMAX* модела приликом предвиђања дужег временског периода унапред.

Сходно чињеници да је *SARIMAX* статички модел, показао је веома добре перформансе приликом краткорочне предикције. То га чини конкурентним у односу на алтернативне моделе машинског учења.

7. ЗАКЉУЧАК

У овом раду представљен је систем за предикцију цена спот инстанци помоћу *SARIMA* модела временске серије. Мотивација проистиче из чињенице да се област рачунарства у облаку рапидно развија и постаје кључни део информатичке инфраструктуре бројних организација. Цена спот инстанци је кључни фактор у управљању трошковима и оптимизацији ресурса у рачунарству у облаку.

Динамичко одређивање цена може закомпликовати процес планирања и алокације ресурса. Стога, тачна предикција цена спот инстанци може бити од велике користи за кориснике облака, омогућавајући им да максимално искористе своје ресурсе и минимизирају трошкове.

У овом раду је за предикцију цена спот инстанци коришћен *SARIMAX* модел за предвиђање временских серија. Извршена су два експеримента која су пружила увид у перформансе модела приликом краткорочне и дугорочне предикције. Евалуација модела извршена је израчунавањем средње апсолутне процентуалне грешке (*MAPE*), која квантификује разлику између стварних и предвиђених вредности и омогућава да се тачно утврди ефикасност модела.

Када је реч о статичким моделима временске серије за предикцију, делује да је *SARIMAX* најбоља опција. Међутим, из јасно наведених недостатака, не мора значити да ће он бити адекватна опција приликом сусретања са различитим типовима проблема. Иако је модел показао добре перформансе приликом краткорочних предвиђања, пре његове употребе у продукцији, требало би истражити да ли је овакво ограничење прихватљиво за корисника.

8. ЛИТЕРАТУРА

- [1] Wenqiang Liu†, Pengwei Wang*†, Ying Meng, Caihui Zhao and Zhaohui Zhang, Cloud spot instance price prediction using kNN regression
- [2] Alkharif S, Lee K, Kim H (2018) Time-series analysis for price prediction of opportunistic Cloud computing resources. 2018 7th international conference on emerging databases. Springer, Singapore, pp 221–229
- [3] Dawei Kong, Shijun Liu, Li Pan, Amazon Spot Instance Price Prediction with GRU Network
- [4] G.E.P. Box, G. Jenkins, "Time Series Analysis, Forecasting and Control", Holden-Day, San Francisco, CA, 1970.
- [5] Vaia I. Kontopoulou 1, Athanasios D. Panagopoulos 2,* , Ioannis Kakkos 1 and George K. Matsopoulos 1, A Review of ARIMA vs. Machine Learning Approaches for Time Series Forecasting in Data Driven Networks
- [6] Paliari, I.; Karanikola, A.; Kotsiantis, S. A comparison of the optimized LSTM, XGBOOST and ARIMA in Time Series forecasting. In Proceedings of the 2021 12th International Conference on Information, Intelligence, Systems & Applications (IISA), Chania Crete, Greece, 12–14 July 2021; pp. 1–7.

Кратка биографија:



Милица Прањкић рођена је 1997. године у Новом Саду. Основне академске студије завршила је 2020. године на Факултету техничких наука, на ком брани и мастер рад 2024. године из области Примењене рачунарске науке и информатика - Интелигентни системи. контакт: m.pranjkic@gmail.com