

РЕЗЮМЕ

на

Stock Price Forecast Based on CNN-BiLSTM-ECA Model

Yu Chen , Ruixin Fang , Ting Liang , Zongyu Sha , Shicheng Li , Yugen Yi , Wei Zhou , and Huilin Song

Received 4 May 2021; Revised 3 June 2021; Accepted 28 June 2021; Published 9 July 2021

Academic Editor: Yi-Zhang Jiang

<https://www.hindawi.com/journals/sp/2021/2446543/>

Прогнозирането на цените на акциите е предизвикателна задача поради високия шум, нелинейността и нестабилността на данните от времеви редове на цената на акциите. За да се осигурят по-добри резултати за прогнозиране на цената на акциите, авторите се предлагат нов модел за прогнозиране на цената на акциите, наречен CNN-BiLSTM-ECA, който комбинира конволюционна невронна мрежа (CNN), двупосочна мрежа за дългосрочна памет (BiLSTM) и механизъм за внимание (AM).

Въведение

Във въведението авторите правят широк обзор на настоящите способности за прогнозиране движенията на фондовите борси, като маркират основните проблеми на съществуващите статистически и ML модели и основанията за въвеждането на DNN индустрията като способ за успешно моделиране на нелинейната връзка във финансовите времеви редове. Обърнато е внимание на преимуществото на LSTM и BiLSTM в моделирането на времеви серии данни. Засегната е темата за възможността за внедряване на Механизъм за внимание (AM) в съвременните архитектури на невронните мрежи прогнозиращи пазарите. Направен е обзор на достиженията в областта. Цитирани са определени научни разработки в тази област.

Вдъхновени от успешните приложения на механизма за дълбоко учене и внимание при анализ на данните за акциите, тази статия е предложен модел за прогнозиране на времеви редове, наречен CNN-BiLSTM-ECA, който хибридна мрежа състояща се от Конволюционна невронна мрежа (CNN) и BiLSTM за прогнозиране на цената на затваряне на данните за акциите. Допълнително е внедрен механизъм за внимание (ECA).

В статията следва по обстоен преглед на основните концепции и методи на машинното обучение, дълбокото машинно обучение, мрежите за дългосрочна памет и механизмите за внимание.

Структура на предложения мрежови модел

Модела претрпява комбиниран с BiLSTM, CNN, модел и лек модул за внимание ECA. Модела е наречен CNN-BiLSTM-ECA.

CNN се използва за извличане на вектори на дълбоки характеристики от входните данни за времеви редове на произход. След това се използва моделът BiLSTM,¹ за да се научат времеви характеристики от новите данни от времеви серии, конструирани от векторите на дълбоките характеристики. Освен това механизъм за внимание, наречен ECA, е допълнително въведен за извличане на по-важни характеристики. И накрая, моделът Dense, състоящ се от няколко напълно свързани слоя, се използва за изпълнение на задачата за прогнозиране. В долната графика е видна архитектурата на конкретната мрежа и етапите на трениране на модела.

¹ Авторите не посочват конкретния начин на реализация на BiDirectional Layers, което оставя отворен въпроса с представителността на прогнозата правена, на трениранировки с "бъдещи" данни. По-долу в статията е упоменато, че се ползват тренировъчни и тестови данни, но няма индикация за валидационни данни.

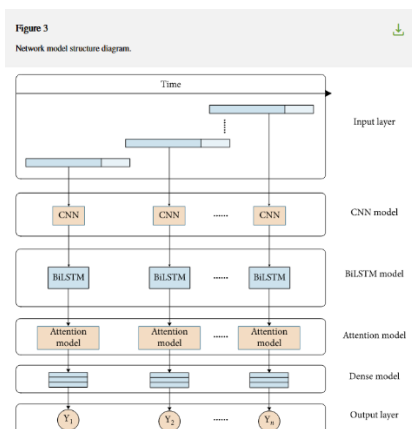


Table 5
Parameters setting of the proposed method.

Description of the parameters		Value
CNN	The filter number of convolution layer in CNN	1
	The kernel_size of the convolution layer in CNN	1
	The activation function in CNN	Relu
	The way of convolution layer padding in CNN	Same
BiLSTM	The number of hidden units of BiLSTM	64
	The activation function of BiLSTM	Relu
Dense	The layers number of the Dense model	3
	The hidden neurons number of the first Dense layer	128
	The hidden neurons number of the second Dense layer	32
	The hidden neurons number of the third Dense layer	1
	The activation function of Dense model	Relu
Other	Time step	10
	Batch size	256
	Epochs	200
	Optimizer	Adam
	Loss function	MSE
	Learning rate	0.001

За да се избегне прекомерния overfitting на модела и намаляване на изчисленията, е конструиран олекотен и ниска сложност модул на вниманието, наречен Effective Channel Attention (ECA). За данните от времеви серии по-големите тегла ще бъдат присвоени за ключовите характеристики и по-малките тегла за нерелевантните характеристики. Ето защо ECA се фокусира върху полезната информация, която подобрява чувствителността на мрежата към основните характеристики.

Експериментален процес

Експериментът се състои от шест части: събиране на данни, предварителна обработка на данни, обучение на модела, запазване на модел, тестване на модела и резултати от прогнозиране.

Експерименталните данни са събрани от NetEase Finance. Включват цена на затваряне, най-висока цена, най-ниска цена, цена на отваряне, цена на затваряне от предходния ден, както и други данни от времеви серии. Ползвани са три сета с данни.

Липсващите стойности на атрибутите са запълнени чрез интерполация. Нормализирането на данните е посредством Min-Max-нормализация.

Конкретната реализация на слоевете на мрежата и размерността им е онагледена във Фигура 9, а описанието на настройките на параметрите на модела е в Таблица 5

Критерии за оценка

Тъй като модела е регресионен се използват три стандартни критерия за оценка на резултатите от прогнозата - средна квадратна грешка (MSE), средноквадратична грешка (RMSE) и средна абсолютна грешка (MAE).

Резултати от експеримента и анализ – стр 11

Авторите анализират параметрите на времевата стъпка в трите набора от данни, включително Shanghai Composite Index, China Mobile и CSI 300. Данните са разделени на тренировъчни и тестови в съотношение 85% :15%. Проведени са сравнителни експерименти, при следните времеви стъпки - 5, 10, 15 и 20. Резултатите са систематизирани и онагледени.

Table 6

Forecast errors at different time steps.

Time step	Shanghai Composite Index			China Unicom			CSI 300		
	MSE	RMSE	MAE	MSE	RMSE	MAE	MSE	RMSE	MAE
5	6431.222	80.195	70.004	0.044	0.209	0.144	3949.252	62.843	43.987
10	1956.036	44.227	28.349	0.028	0.167	0.103	3434.408	58.604	39.111
15	10162.775	100.811	91.489	0.038	0.196	0.177	5013.827	70.808	51.035
20	4711.037	68.637	58.377	0.035	0.186	0.123	6409.845	80.062	57.951

Авторите достигат до извода, че моделът за прогнозиране може да постигне най-добра производителност, когато времевата стъпка е настроена на 10. Направен е сравнителен анализ на ефективността на модела с други архитектури.

Table 7

Forecast errors of different network models.

Model	Shanghai Composite Index			China Unicom			CSI 300		
	MSE	RMSE	MAE	MSE	RMSE	MAE	MSE	RMSE	MAE
CNN	8447.149	91.908	79.914	0.037	0.193	0.134	6218.092	78.855	63.981
LSTM	4222.102	64.978	50.585	0.036	0.189	0.128	5809.153	76.218	58.679
BiLSTM	2603.726	51.027	35.419	0.035	0.187	0.132	5091.610	71.356	52.119
CNN-LSTM	2902.648	53.876	39.502	0.030	0.174	0.110	4905.472	70.039	52.457
CNN-BiLSTM	2321.235	48.179	32.289	0.029	0.170	0.110	4643.541	68.144	51.143
BiLSTM-ECA	2184.278	46.736	32.737	0.039	0.198	0.142	4161.203	64.507	46.453
CNN-LSTM-ECA	2200.705	46.911	32.353	0.032	0.180	0.127	4568.808	67.593	51.061
CNN-BiLSTM-ECA	1956.036	44.227	28.349	0.028	0.167	0.103	3434.408	58.604	39.111

Резултатите показват, че BiLSTM и LSTM са по-добри от модела на CNN. Производителността на BiLSTM е по-добра от LSTM поради факта, че моделът BiLSTM може да използва последващата информация от времеви редове на цените на акциите. Второ, въвеждането на модела на CNN за

намаляване на шума и улавяне на нелинейна структура на запасните данни, методите CNN-LSTM и CNN-BiLSTM могат да превъзхождат CNN, LSTM и BiLSTM. И накрая, чрез интегриране на модела ECA в LSTM и BiLSTM за избор на важни характеристики и ключова информация, методите BiLSTM-ECA, CNN-LSTM-ECA и CNN-BiLSTM-ECA ²могат да постигнат по-добра производителност от другите методи.

Заклучение

Този документ предлага мрежов модел за прогнозиране на времеви серии на цената на акциите (CNN-BiLSTM-ECA), който взема цената на затваряне на акциите, най-високата цена, най-ниската цена, цената на отваряне, цената на затваряне от предходния ден, промяната, както и други данни от времеви серии като вход за прогнозиране на цената на затваряне на акциите на следващия ден. Предложеният мрежов модел комбинира мрежови модели на CNN и BiLSTM. Първо, CNN се използва за ефективно извличане на дълбоките характеристики на входните данни. Второ, векторите на характеристиките се конструират във времеви серии като входове към мрежата BiLSTM за обучение и прогнозиране. В същото време моделът на вниманието на ECA се въвежда в модела, за да се подобри важността на функциите за обучение. Експерименталните резултати показват, че предложеният модел има най-висока точност на прогнозиране и най-добра производителност.

Авторите имат идея за увеличаване на характеристиките за обучение на модела с допълнителни данни за сентимент на пазара, данни от медии и социални мрежи, така че да се подобри ефективността на прогнозиране на модела.

² Прави впечатление, че на графиките реални/прогнозни данни ясно личи отместване, което показва неточност в прогнозата - прогнозата изостава от реалната ситуация на пазара. Авторите не са коментирали това.