



پیش بینی داده های سری زمانی با استفاده از شبکه های عصبی عمیق (داده های مصرفی برق – داده های فرودگاه)

حسین حسین زاده^۱ – رضا باقرپور^۲

دانشکده مهندسی کامپیوتر

دانشگاه علم و صنعت

[1. h.hosseinzade1994@gmail.com](mailto:h.hosseinzade1994@gmail.com) – [2. rbagherpour72@yahoo.com](mailto:rbagherpour72@yahoo.com)

چکیده

بررسی گذشته داده ها به منظور شناسایی و تعیین ماهیت یک پدیده، قدم اول در یک مدل سازی اصولی است که با استفاده از آن می توان در گام بعدی به یک پیش بینی صحیح و قابل اتکا دست یافت. پیش بینی بار روزانه در شرکت های توزیع که به منظور ارائه به شرکت مدیریت شبکه و همچنین به منظور اتخاذ روشی مناسب در قبال مدیریت بار می باشد امری مهم در شبکه های توزیع محسوب می گردد، در این حوزه پیش بینی قیمت برق نیز برای مشتریان و همچنین شرکت های توزیع از اهمیت بالایی برخوردار است، از این رو در این مقاله ابتدا به بررسی مسئله پیش بینی بار مصرفی برق مشتریان خواهیم پرداخت و برای ارزیابی مدل ارائه شده، این مدل را برای پیش بینی قیمت برق با داده های متفاوت به کار خواهیم برد. در ادامه کار به منظور جامعیت بخشیدن به مسئله پیش بینی نوع دیگری از داده های سری زمانی به نام داده های سری زمانی چند متغیره را مورد پیش بینی و ارزیابی قرار خواهیم داد این نوع داده های سری زمانی چند متغیره در حوزه هایی که می توان ویژگی های بیشتری را برای داده ها در نظر گرفت از اهمیت قابل توجهی برخوردار می باشند.

۱ مقدمه

سری زمانی توالی از مشاهدات بر روی یک متغیر است که در نقاط متوالی از زمان یا دوره های متوالی در زمان محاسبه می شود. این محاسبه ممکن است هر ساعت، هر روز، هر هفته، هر ماه یا هر سال و یا در هر فاصله زمانی دیگری انجام شود. الگوی داده ها در درک اینکه سری زمانی در گذشته چگونه رفتار کرده بسیار مهم است. اگر انتظار رود که چنین رفتاری در آینده هم ادامه پیدا کند، می توان از الگوهای گذشته برای راهنمایی بهتر در انتخاب روش پیش بینی مناسب برای آینده هم استفاده کرد. هدف آنالیز سری های زمانی کشف الگویی در داده های گذشته یا سری های زمانی و سپس پیش بینی الگویی برای آینده است؛ این پیش بینی تنها بر مقادیر گذشته متغیر و/ یا خطاهای پیش بینی های گذشته تکیه دارد.

در سالهای اخیر با ثبت منظم داده ها و به وجود آمدن بانک های مختلف اطلاعاتی در سراسر دنیا تمایل به استفاده از سری های زمانی افزایش یافته است. سری های زمانی که زیرشاخه ای از آمار و احتمالات میباشد، مجموعه ای از مشاهدات است که برحسب زمان مرتب شده است. [۱] امروزه با آگاه تر شدن پژوهشگران از کاربردهای سری های زمانی به طور وسیع در خدمات بهداشتی و درمانی، صنایع، انرژی، حمل و نقل و بسیاری از حوزه های دیگر مورد استفاده قرار می گیرد. [۲]

۱-۱ داده های سری زمانی تک متغیره

داده های سری زمانی تک متغیره از آن جهت که تنها دارای یک ویژگی وابسته به زمان برای پیش بینی مقادیر آینده هستند به عنوان داده های سری زمانی تک متغیره تلقی می شوند. این نوع داده ها به خاطر تک متغیره بودن از درجه سخت تری برای پیش بینی قرار دارند.

پیش بینی بار روزانه در شرکت های توزیع که به منظور ارائه به شرکت مدیریت شبکه و همچنین به منظور اتخاذ روش مناسب مدیریت بار می باشد جدا ناپذیر در شبکه های توزیع محسوب می گردد و نیز پیش بینی دقیق بار قیمت تولید برق را رد شبکه قدرت کاهش خواهد داد و در بحث بهره برداری و قابلیت اطمینان شبکه قدرت موثر و امری ضروری می باشد و لذا با توجه به تاثیرگذاری الگوهای بار از عوامل مختلفی مانند آب و هوایی، اقتصادی، اجتماعی و غیره پیش بینی دقیق امر دشواری می باشد. به همین دلیل در سال های اخیر با استفاده از روش های مختلف به مسائل پیش بینی بار پرداخته اند.

داده هایی که برای پیش بینی در این حوزه استفاده می شوند اغلب داده های سری زمانی هستند که از مقدار بار مصرفی در روز یا زمان خاصی را نشان می دهند. [۳] برای آموزش مدل ها و ارزیابی نتایج در داده های مصرفی انرژی الکتریکی، که در فواصل زمانی ۵ دقیقه ای ثبت شده اند، داده ها در همه ی مدل ها به دو قسمت داده های آموزش و تست تقسیم شده اند که داده های آموزش شامل ۹۰۰۰ نمونه از این نوع داده ها و داده های تست نیز شامل ۱۲۰ نمونه برای ارزیابی پیش بینی بار می باشد.

از دیگر مسائلی که در این حوزه وجود دارد مسئله پیش بینی قیمت برق در ساعات مختلف روز و یا دوره های بلند مدت می باشد، این نوع پیش بینی به مشتریان در مصرف مناسب انرژی در ساعات مناسب روز جهت کاستن از هزینه پرداختی بابت مصرف انرژی برق کمک می کند و از یک رو به شرکت های توزیع برق این امکان را می دهد که با استفاده از الگوی قیمت گذاری مناسب مدیریت انرژی را بهتر اداره کنند. این نوع داده ها نیز در برخی از شرکت های توزیع در دسترس می باشند [۴] و در این مقاله به عنوان داده های سری زمانی مورد ارزیابی قرار می گیرند. در داده های قیمت برق داده های آموزش شامل ۱۷۵۲۱ نمونه از این نوع داده ها و داده های تست نیز شامل ۴۸ نمونه برای ارزیابی پیش بینی قیمت می باشد.

۱-۲ داده های سری زمانی چند متغیره

داده های سری زمانی چند متغیره داده هایی هستند که دارای ویژگی های بیشتری می باشند و از این رو وابسته به چندین پارامتر مختلف جهت پیش بینی می باشند.

یکی از حوزه هایی که پیش بینی داده های سری زمانی می تواند در آن اهمیت داشته باشد مسئله تعداد مسافران روزانه فرودگاه های هواپیمایی است که می توان بر اساس آن تعداد پرواز ها، خدمات جانبی و مسائل بسیاری را به ازای پیش بینی های انجام شده از قبل برنامه ریزی کرد. داده هایی که در این بخش مورد استفاده قرار گرفته است از سایت خرید بلیط هواپیمایی علی بابا تهیه شده است. که این مجموعه داده ها متشکل از نمونه های متعدد ثبت شده ی بلیط های خریداری شده ی روزانه به مدت زمان ۲ سال می باشد. داده های به کار برده شامل اطلاعات خرید مربوط به ۲ سال می باشد که شامل ۷۳۰ نمونه می باشد، که تعداد ۶۰۰ روز به عنوان آموزش در نظر گرفته شده است و مابقی به عنوان تست شبکه مورد ارزیابی قرار می گیرد.

۲ کارهای مرتبط

پیش بینی مصرف انرژی الکتریکی در حوزه های بسیاری کاربرد دارد. علاوه بر بازارهای برق در مدیریت سمت تقاضا و کنترل نرخ افت و خیز برق نیز استفاده می شود. روش های پیش بینی معمولی نظیر SVR(support vector regression)، شبکه های عصبی مصنوعی (ANNs) روش های منطق فازی [۵] و متد های آنالیز سری زمانی مثل ARIMA [۶] و مدل های خاکستری (GMs). در سال ۲۰۰۷ Ediger و Akar [۷] مدل ARIMA را در شروع برای پیش بینی انرژی مصرفی تا سال ۲۰۲ را در ترکیه به کار بردند. Ogcı [۸] شبکه عصبی مصنوعی و مدل SVM را برای پیش بینی مصرف انرژی به کار برد، در این مطالعه برای اندازه گیری میزان دقت مدل از نرخ MAPE(mean absolute percentage error) استفاده کردند. که مدل SVR، ۶ درصد کارایی مناسب تری نسبت به ANNs نشان داد.

شبکه های عصبی مصنوعی عمیق تکنیک های جدید یادگیری ماشین هستند که مورد استقبال قرار گرفته اند. این تکنیک با داده های بزرگ سرو کار دارد که دقت پیش بینی و کلاس بندی خوبی را در مسائلی همچون پیش بینی ضریب سهام [۹ و ۱۰]، پیش بینی سرعت باد [۱۱]، پیش بینی تابش خورشید [۱۲ و ۱۳] و غیره نشان داده اند. در سال های اخیر با پیشرفت های سریع در حوزه شبکه های هوشمند، شبکه های عصبی عمیق به صورت گسترده ای در جهت پیش بینی قیمت انرژی و مقدار مصرف انرژی مورد استفاده قرار گرفته اند. Ryu [۱۴] شبکه عمیق عصبی را با داده های مصرفی خانه ای که در طی سال ۲۰۱۶ ثبت شده بود را آموزش داد، نتیجه این کار نشان می دهد که شبکه عمیق دقت پیش بینی بهتری را در مقایسه با شبکه عصبی shallow، مدل DSHW(double seasonal Holt-Winters) و مدل ARIMA دارد.

۳ مدل پیشنهاد شده

در داده ای سری زمانی تک متغیره که تنها یک مقدار وابسته به زمان وجود دارد، شبکه های عصبی عمیق در حالت عادی نتایج خوبی را نسبت به دیگر مدل ها نشان نمی دهند. برای بهبود مدل شبکه های عمیق در برخورد با این نوع داده ها می توان از لایه های کانولوشن تک بعدی جهت استخراج ویژگی های بیشتر در ابتدای شبکه استفاده کرد. لایه ی کانولوشن یک بعدی (1-D convolution) یکی

از ویژگی های مدل پیشنهادی می باشد که برای پردازش داده های سری زمانی تک متغیره مناسب است. از این رو در این بخش با استفاده از لایه کانولوشن که از یک کرنل $k*1$ استفاده می کند می توانیم داده های تک متغیره اصلی را به داده هایی با ویژگی های چند بعدی تغییر دهیم. استفاده از شبکه های LSTM یکی دیگر از ویژگی های این شبکه می باشد. LSTM در پیش بینی برای مسائل دارای توالی بسیار قدرتمند عمل می کند، زیرا قادر به ذخیره سازی اطلاعات گذشته است. از این رو مدل مورد نظر برای رسیدن به یک نتیجه مناسب استفاده از شبکه ای بود که تلفیقی از لایه های کانولوشن و LSTM باشد، که بتواند در برابر داده های تک متغیره عملکرد بهتری داشته باشد.

با توجه به مسائل در مرحله اول مدل پیشنهادی را برای پیش بینی مصرف انرژی استفاده میکنیم، سپس مدل پیشنهادی را تعمیم داده و برای مسئله پیش بینی تعداد مسافران و قیمت برق که داده های آن ارتباطی به داده های مصرفی ندارند، به کار می بریم.

۴ نتایج

یک عامل مهم در تعیین روش پیش بینی مورد استفاده دقت روش است. سه عامل دقت پیش بینی مورد بحث قرار می گیرند (MAE) mean absolute error ، (MSE) mean squared error و (MAPE) mean absolute percentage error. هر یک از این معیارها برای تعیین اینکه یک روش پیش بینی تا چه حد می تواند داده های سری های زمانی را که اکنون در دست هستند مجددا تولید کند، طراحی شده اند. با انتخاب روشی که بهترین دقت را برای داده های شناخته شده فعلی دارد، امیدواریم احتمال به دست آوردن پیش بینی های بهتر برای دوره های زمانی آتی را افزایش دهیم.

۴-۱ داده های انرژی مصرفی برق

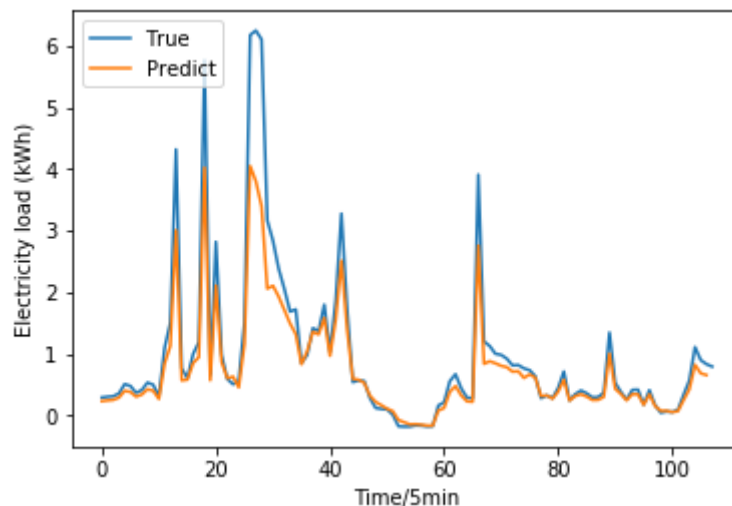
۴-۱-۱ مدل پایه (SVR)

اولین مدل مورد بررسی در پیش بینی داده های سری زمانی که مناسب این نوع از داده ها نیز می باشد، مدل SVR (Support Vector Regression) است که پس از انتخاب پارامترهای مناسب آموزش، نتایج زیر از داده های استفاده شده در مرحله تست حاصل شد.

مقداردهی پارامترهای مدل :

- ❖ kernel type = rbf
- ❖ $C = 1000$ (Penalty parameter C of the error term)
- ❖ $\gamma = 0.0001$ (Kernel coefficient)
- ❖ $\epsilon = 0.1$

نمودار مقایسه مقدار بار واقعی مصرف شده با مقدار بار پیش بینی شده با مدل SVR:



شکل (۱) بار حقیقی و بار پیش بینی شده در مدل SVR

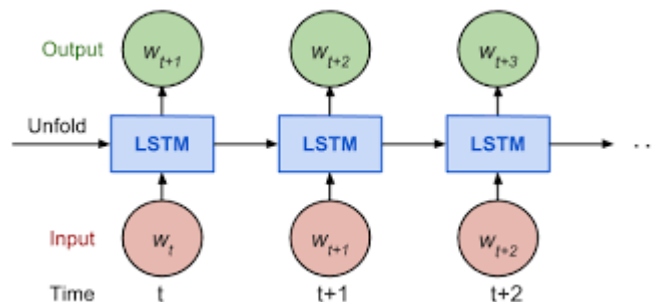
نتایج معیارهای ارزیابی استفاده شده برای مدل :

- RMSE: 1.055
- MAPE: 20.909
- MAE: 0.500

۴-۱-۲ شبکه LSTM

LSTM در پیش‌بینی برای مسائل دارای توالی بسیار قدرتمند عمل می‌کند، زیرا قادر به ذخیره‌سازی اطلاعات گذشته است. از این رو اولین مدل یادگیری عمیق که از آن برای پیش‌بینی بار استفاده شد مدل LSTM بود.

ساختار شبکه LSTM :



• مدل شبکه:

مدل اولیه مورد نظر عبارتند از ۲ لایه شبکه LSTM هر کدام با ۵۰ نورون و یک لایه Dense برای خروجی شبکه به اندازه ۱.

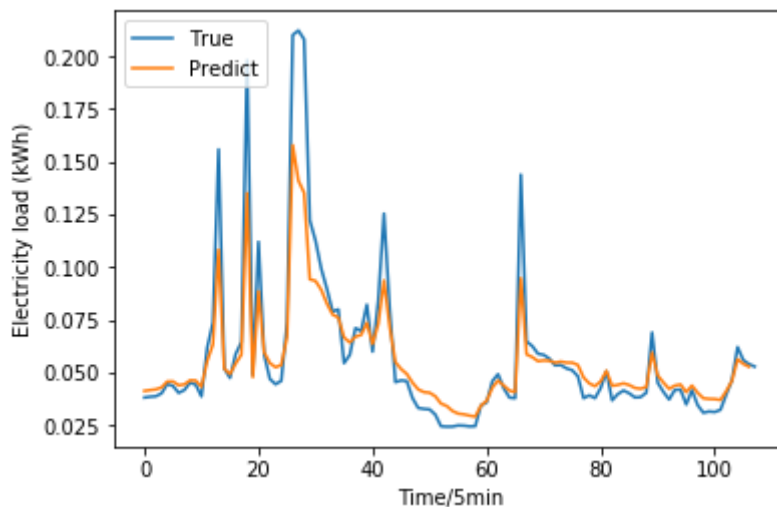
Layer (type)	Output Shape	Param #
lstm_16 (LSTM)	(None, 12, 50)	10400
lstm_17 (LSTM)	(None, 50)	20200
dense_7 (Dense)	(None, 1)	51
Total params: 30,651		
Trainable params: 30,651		
Non-trainable params: 0		

شکل ۲) مدل LSTM + DENSE

- ورودی شبکه :

LSTM، از داده‌ها انتظار دارد تا در قالب مشخصی باشند، که معمولاً آرایه‌های سه بعدی است. کار با ساخت داده‌ها در ۱۲ گام زمانی (Time Step) و تبدیل آن‌ها به یک آرایه انجام می‌شود. سپس، داده‌ها به یک آرایه سه بعدی شامل نمونه‌های آموزش، تعداد ۱۲ گام زمانی و یک ویژگی در هر گام تبدیل می‌شوند. در نتیجه شکل ورودی برای لایه LSTM $\text{input_shape} = (12, 1)$ خواهد بود.

نمودار مقایسه مقدار بار واقعی مصرف شده با مقدار بار پیش بینی شده با مدل LSTM :



شکل ۳) بار حقیقی و بار پیش بینی شده در مدل LSTM

نتایج معیارهای ارزیابی استفاده شده برای مدل :

- RMSE: 0.029
- MAPE: 12.477
- MAE: 0.015

۳-۱-۴ شبکه CNN+LSTM

در داده ای سری زمانی تک متغیره که تنها یک مقدار وابسته به زمان وجود دارد، شبکه های عصبی عمیق در حالت عادی نتایج خوبی را نسبت به دیگر مدل ها نشان نمی دهند. برای بهبود مدل شبکه های عمیق در برخورد با این نوع داده ها می توان از لایه های کانولوشن تک بعدی جهت استخراج ویژگی های بیشتر در ابتدای شبکه استفاده کرد. لایه ی کانولوشن یک بعدی (1-D convolution) یکی از ویژگی های مدل پیشنهادی می باشد که برای پردازش داده های سری زمانی تک متغیره مناسب است. از این رو در این بخش با استفاده از لایه کانولوشن که از یک کرنل $k \times 1$ استفاده می کند می توانیم داده های تک متغیره اصلی را به داده هایی با ویژگی های چند بعدی تغییر دهیم.

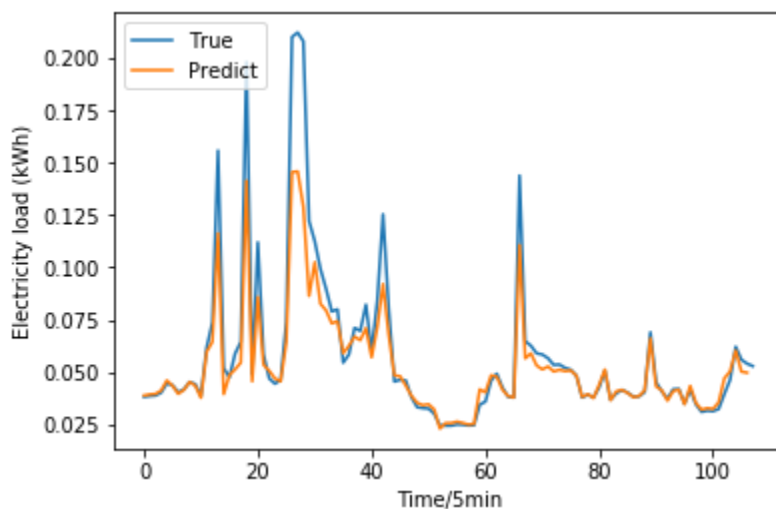
• مدل

استفاده از لایه ی کانولوشن در ابتدای شبکه، و قرار دادن لایه LSTM و یک لایه Dense با یک خروجی در انتهای شبکه.

Layer (type)	Output Shape	Param #
conv1d_5 (Conv1D)	(None, None, 32)	192
conv1d_6 (Conv1D)	(None, None, 64)	6208
lstm_15 (LSTM)	(None, 100)	66000
dense_6 (Dense)	(None, 1)	101
Total params: 72,501		
Trainable params: 72,501		
Non-trainable params: 0		

شکل ۴) شبکه CONV+LSTM

نمودار مقایسه مقدار بار واقعی مصرف شده با مقدار بار پیش بینی شده با مدل CNN+LSTM :



شکل ۵) بار حقیقی و بار پیش بینی شده در مدل CNN+LSTM

نتایج معیارهای ارزیابی استفاده شده برای مدل :

- RMSE: 0.030
- MAPE: 7.396
- MAE: 0.014

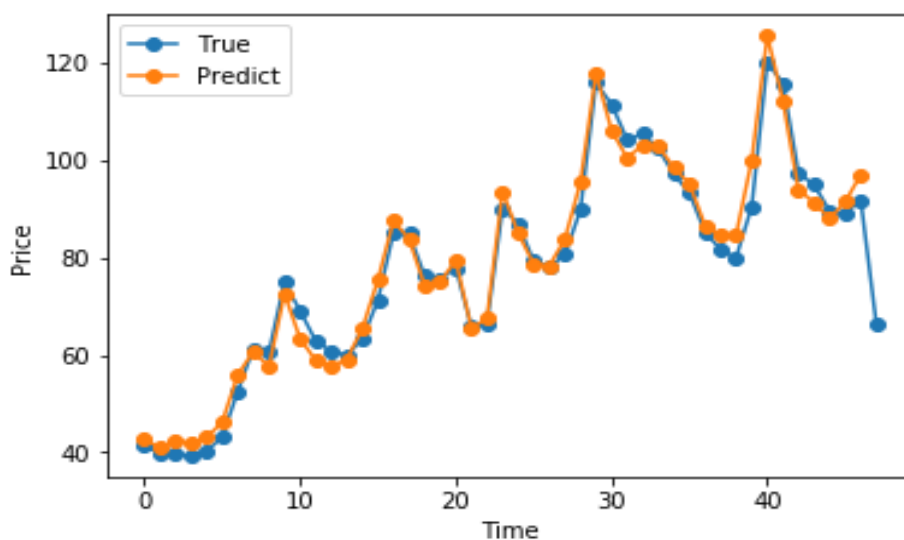
error model	RMSE	MAPE	MAE
SVR	1.055	20.909	0.500
LSTM	0.29	12.477	0.015
CNN+LSTM	0.030	7.396	0.014

جدول ۱) مقایسه خطای مدل های مختلف

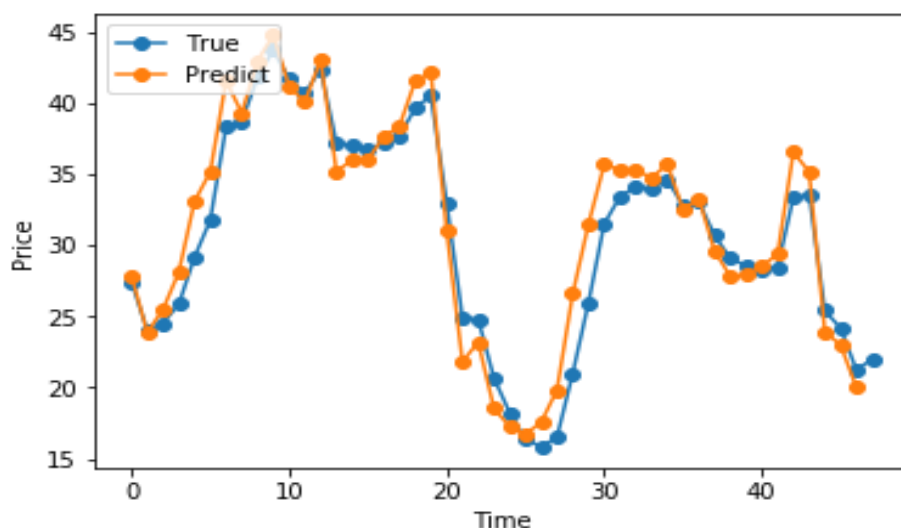
۲-۴ داده های قیمت انرژی مصرفی (تعمیم شبکه استفاده شده در مرحله قبل)

۱-۲-۴ شبکه CNN+LSTM برای داده های قیمت برق

در مقاله ای به عنوان (۱۵) Electric Power Price Forecasting Using Data Mining Techniques که در سال ۲۰۱۷ منتشر شده است، از روش کلاسترینگ و مدل ARIMA برای پیش بینی قیمت برق مصرفی استفاده شده است در این کار ابتدا داده ها با استفاده از دسته بندی k-means بر اساس روزهای هفته به سه دسته کلاس بندی می شوند. سپس این داده های کلاس بندی شده با تکنیک KNN بر اساس ماه های سال به دو الگو کلاس بندی می شوند و در انتها با استفاده از مدل ARIMA مورد پیش بینی قرار می گیرند. در این قسمت مدل شبکه عمیقی را که برای پیش بینی مقدار مصرف انرژی به کار برده بودیم، در این بخش برای داده های قیمت برق که هیچ ارتباطی به داده های مصرفی برق ندارند تعمیم می دهیم با این کار یک مدل مناسب برای پیش بینی داده های مختلف بدست آمده است.



شکل ۶) پیش بینی مدل پیشنهادی برای روزهای ۵ و ۶ February



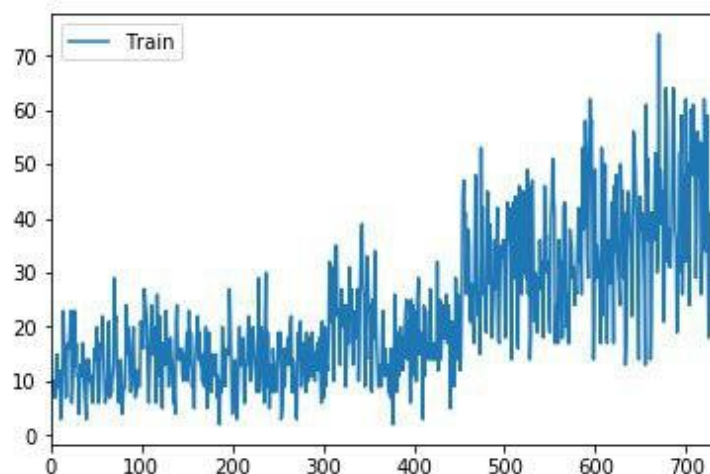
شکل ۷) پیش بینی مدل پیشنهادی برای روزهای ۹ و ۱۰ February

در مقاله ذکر شده مقدار MAPE که از پیش بینی مدل مقاله برای روزهای ۵-۶ February و ۹-۱۰ February بدست آمده است به ترتیب برابر $4/16$ و $7/24$ می باشد، در حالی که با استفاده از شبکه عصبی عمیق تعمیم داده شده این مقادیر به ترتیب به $3/57$ و $5/7$ کاهش یافتند. و این دقت بالای شبکه عمیق نسبت به مدل استفاده شده در مقاله را نشان می دهد.

۳-۴ داده های خرید بلیط هواپیما

۱-۳-۴ پیش بینی با استفاده از مدل پیشنهادی

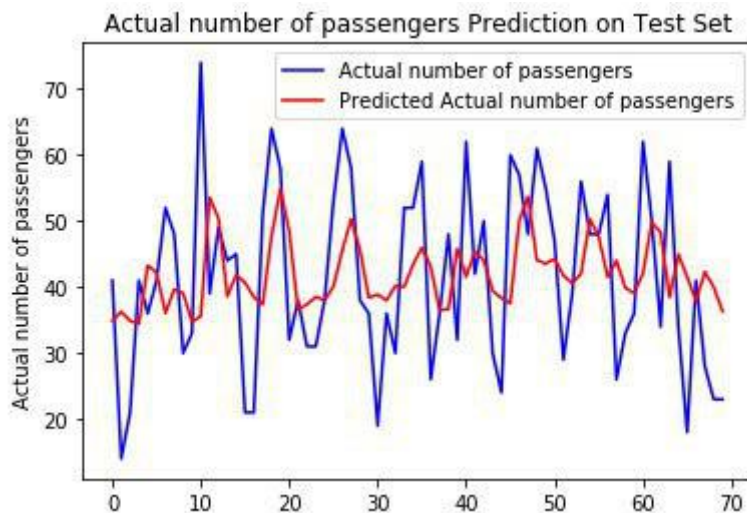
همانطور که در شکل زیر مشاهده می کنید این نوع داده ها به دلایلی از جمله تغییر رجوع مشتری ها در طی این مدت، از توزیع خوبی برخوردار نیستند و این امر باعث شده است که پیش بینی این داده ها از سختی بالایی نسبت به داده های نرمال برخوردار باشد. از این رو با استفاده از مدل پیشنهادی که برای داده های برق در نظر گرفته بودیم با توجه به مناسب نبودن داده ها نتیجه قابل قبولی از این پیش بینی حاصل نشد و انتظاری هم از این امر نبود.



شکل ۸) داده های سری زمانی مربوط به تعداد مسافرین روزهای مختلف سال ۹۵-۹۶

Layer (type)	Output Shape	Param #
conv1d_7 (Conv1D)	(None, None, 32)	192
conv1d_8 (Conv1D)	(None, None, 32)	3104
lstm_17 (LSTM)	(None, None, 50)	16600
dropout_17 (Dropout)	(None, None, 50)	0
lstm_18 (LSTM)	(None, None, 50)	20200
dropout_18 (Dropout)	(None, None, 50)	0
lstm_19 (LSTM)	(None, None, 50)	20200
dropout_19 (Dropout)	(None, None, 50)	0
lstm_20 (LSTM)	(None, 50)	20200
dropout_20 (Dropout)	(None, 50)	0
dense_5 (Dense)	(None, 1)	51
Total params: 80,547		
Trainable params: 80,547		
Non-trainable params: 0		

شکل ۹) مدل استفاده شده برای پیش بینی تعداد مسافرین با استفاده از مدل پیشنهادی



شکل ۱۰) خروجی حاصل از پیش بینی تعداد مسافری با شبکه پیشنهادی

۴-۳-۲ تبدیل داده ی سری زمانی تک متغیره به سری زمانی چند متغیره

برای حل چنین مسئله ای با همچنین داده ای از تکنیک استخراج ویژگی استفاده شد که باعث شد داده ها از حالت سری زمانی تک متغیره به سری زمانی چند متغیره تبدیل شود و این امر باعث بهبود در پیش بینی شد.

Log_Date	Count	Holiday	day_of_month	day_of_week(N)	month_of_year	Future_Holiday_NUM(7)	Future_Holiday_NUM(14)
1395/01/02	44	1	2	3	1	4	6
1395/01/03	38	1	3	4	1	3	5
1395/01/04	49	1	4	5	1	2	4
1395/01/05	81	1	5	6	1	2	4
1395/01/06	40	1	6	7	1	2	4

شکل ۱۱) استخراج ویژگی های مورد نیاز در این مسئله از روی تاریخ پرواز

این ویژگی ها عبارتند از:

Holiday: روزهای تعطیل رسمی سال و همچنین روزهای پنجشنبه و جمعه هر هفته

day_of_week: روزهای هر هفته (این ویژگی با توجه به معنی دار نبودن اختلاف در مقادیر آن به one-hot تبدیل شد)

day_of_month: روزهای ماه (one-hot)

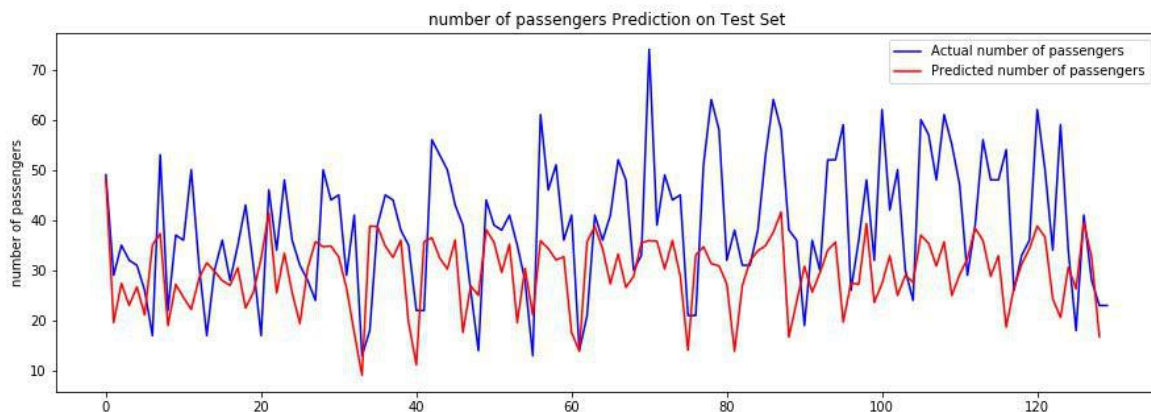
month_of_year: ماه های سال که به دلیل تاثیرگذاری فصول مختلف از اهمیت خوبی برخوردار است. (one-hot)

future_holiday_num(7): اشاره به تعداد روزهای تعطیل در ۷ روز آینده دارد.

future_holiday_num(14): اشاره به تعداد روزهای تعطیل در ۱۴ روز آینده دارد.

۳-۳-۴ پیش بینی سری زمانی چند متغیره با مدل Random Forest

مدل پایه ای که برای این داده های در نظر گرفته شد مدل قدرتمند Randomo Forest بود تا با استفاده از ویژگی های متعدد بتواند نتیجه خوبی را پیش بینی کند. شکل زیر حاصل پیش بینی ۱۳۰ روز منتهی به آخر سال را نشان می دهد.



شکل ۱۳) پیش بینی حاصل از مدل Random Forest

برای اندازه گیری دقت از معیار RMSE استفاده کرده ایم که مقدار آن برای مدل Random Forest ۰/۲ می باشد.

۴-۳-۴ استفاده از شبکه LSTM برای سری زمانی چند متغیره

برای استفاده از شبکه LSTM برای داده های سری زمانی لازم است که داده ها به مسئله یادگیری با نظارت تبدیل شوند یعنی به ازای هر سری از ورودی برای آن خروجی در نظر گرفته شود.

برای مثال اگر داده های سری زمانی تنها شامل اعداد ۱ تا ۹ باشند می توان با تبدیل این داده ها به مسئله یادگیری با نظارت مدل های خود را بر روی آن اعمال کرد. به این صورت که نمونه بعدی به عنوان خروجی آن نمونه محسوب می شود.

		var1(t-1)	var1(t)
1			
2	1	0.0	1
3	2	1.0	2
4	3	2.0	3
5	4	3.0	4
6	5	4.0	5
7	6	5.0	6
8	7	6.0	7
9	8	7.0	8
10	9	8.0	9

شکل ۱۴) داده های سری زمانی تبدیل شده به داده های مسئله ی یادگیری با نظارت (صرفاً به عنوان مثال)

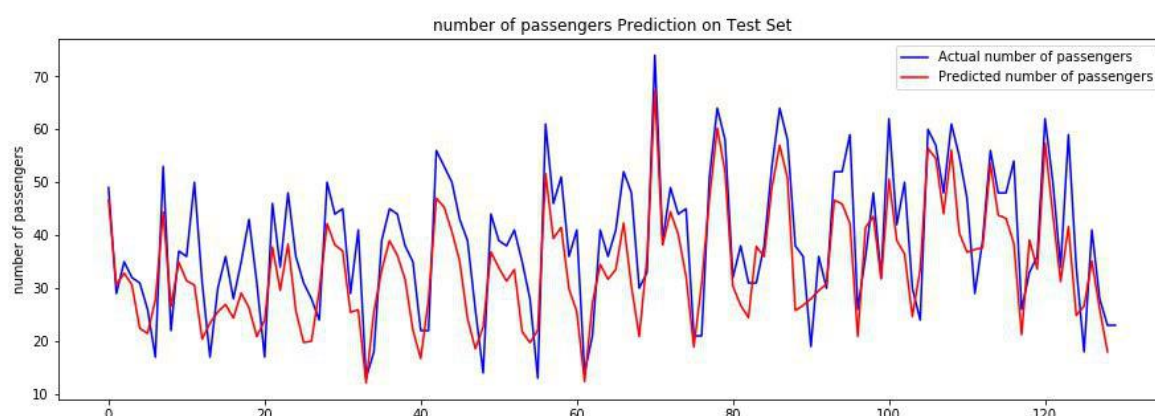
پس از تبدیل داده ها به یک مسئله ی یادگیری با نظارت با استفاده از شبکه عصبی عمیق که متشکل از چندین لایه LSTM می باشد می توان یادگیری را انجام داد.

• مدل

Layer (type)	Output Shape	Param #
lstm_1 (LSTM)	(None, 1, 50)	19000
dropout_1 (Dropout)	(None, 1, 50)	0
lstm_2 (LSTM)	(None, 1, 50)	20200
dropout_2 (Dropout)	(None, 1, 50)	0
lstm_3 (LSTM)	(None, 1, 50)	20200
dropout_3 (Dropout)	(None, 1, 50)	0
lstm_4 (LSTM)	(None, 50)	20200
dropout_4 (Dropout)	(None, 50)	0
dense_1 (Dense)	(None, 1)	51
Total params: 79,651		
Trainable params: 79,651		
Non-trainable params: 0		

شکل ۱۵) مدل پیشنهادی برای داده های فرودگاه

یادگیری این مدل نیز همانند مدل قبلی با استفاده از ۶۰۰ داده انجام می شود و با ۱۳۰ داده نیز مدل مورد ارزیابی قرار خواهد گرفت.



شکل ۱۵) پیش بینی حاصل از مدل LSTM

معیار RMSE برای این شبکه ۰/۱۴ را نشان می دهد که در مقایسه با مدل Random Forest از عملکرد بهتری برخوردار است.

۵ تحلیل

در داده های سری زمانی که به صورت تک متغیره مورد بررسی قرار گرفتند، دو دلیل وجود کافی داده ها از نظر کمی و همچنین کیفیت داده ها از نظر stable بودن، این امکان را فراهم آورد که بتوان با استفاده از مدل های مناسب پیش بینی خوبی را انجام داد. مدل های پایه ای یادگیری ماشین در مقابل چنین داده هایی که تنها یک متغیر وابسته به زمان را در اختیار دارد تا حدودی جواب قابل قبولی را نشان دادند، در شبکه های عمیق مدل مناسبی که این گونه پیش بینی ها را انجام دهد استفاده از شبکه LSTM بود، اما در شبکه های عمیق می توان با استفاده از خاصیت کانولوشن تک بعدی ویژگی های ای نوع داده ها را افزایش داد که این کار تاثیر مستقیمی را در دقت پیش بینی نشان داد و حتی به عنوان مدلی تعمیم یافته از آن در داده های متفاوت استفاده شد.

در داده های سری زمانی که از نظر تعداد نمونه دارای محدودیت هستند و همچنین از نظر کیفی در حد مطلوبی قرار ندارند، استفاده مستقیم از داده های تک متغیره، به ازای پیش بینی داده ها، نتایج خوبی

را در خروجی مدل ها شاهد نخواهیم بود. از این رو در مقابل چنین داده هایی از استخراج ویژگی ها استفاده می کنیم تا بتوانیم با توجه به محدودیت داده ها، با کمک گرفتن از متغیرهای بیشتر بتوانیم پیش بینی مناسبی را در خروجی مدل ها مشاهده کنیم. با در نظر گرفتن چنین اعمالی روی داده ها، شاهد ایجاد نوع جدیدی از داده های سری زمانی خواهیم شد که در بهبود پیش بینی از اهمیت چشم گیری برخوردار است به خصوص برای استفاده در شبکه های LSTM که دقت پیشی بینی مسائل سری زمانی چند متغیره را با بهبود محسوسی به دنبال دارد.

1. Koopman SJ. Time series analysis by state space methods: OUP Oxford; 2012.. 93-113.
2. Hasanzadeh J, Amiresmaili M, Moosazadeh M, Najafi F, Moradinazar M. Implementing a Weather-Based Early Warning System to Prevent Traffic Accidents Fatalities. World Applied Sciences Journal. 2013; 24: 113-7.
3. Kelly, J.; Knottenbelt, W. The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes. Sci. Data 2015
4. The NYISO Website [Online]. Available: <http://www.nyiso.com>
5. Karnik, N.N.; Mendel, J.M. Applications of type-2 fuzzy logic systems to forecasting of time-series. Inform. Sci. 1999, 120, 89–111.
6. Conejo, A.J.; Plazas, M.A.; Espinola, R.; Molina, A.B. Day-ahead electricity price forecasting using the wavelet transform and ARIMA models. IEEE Trans. Power Syst. 2005, 20, 1035–1042.
7. Ediger, V. ,S.; Akar, S. ARIMA forecasting of primary energy demand by fuel in Turkey. Energy Policy 2007, 35, 1701–1708.
8. O ̇gcu, G.; Demirel, O.F.; Zaim, S. Forecasting electricity consumption with neural networks and support vector regression. Procedia-Soc. Behav. Sci. 2012, 58, 1576–1585.
9. Armano, G.; Marchesi, M.; Murru, A. A hybrid genetic-neural architecture for stock indexes forecasting. Inform. Sci. 2005, 170, 3–33. [CrossRef]
10. Rather, A.M.; Agarwal, A.; Sastry, V. Recurrent neural network and a hybrid model for prediction of stock returns. Expert Syst. Appl. 2015, 42, 3234–3241.
11. [CrossRef] 37. Khodayar, M.; Kaynak, O.; Khodayar, M.E. Rough deep neural architecture for short-term wind speed forecasting. IEEE Trans. Ind. Inf. 2017, 13, 2770–2779.
12. Voyant, C.; Notton, G.; Kalogirou, S.; Nivet, M.L.; Paoli, C.; Motte, F.; Fouilloy, A. Machine learning methods for solar radiation forecasting: A review. Renew. Energy 2017, 105, 569–582. [CrossRef] Energies 2018, 11, 3089 15 of 15
13. Alzahrani, A.; Shamsi, P.; Dagli, C.; Ferdowsi, M. Solar irradiance forecasting using deep neural networks. Procedia Comput. Sci. 2017, 114, 304–313
14. Ryu, S.; Noh, J.; Kim, H. Deep neural network based demand side short term load forecasting. Energies 2016, 10, 3.