

بسمه تعالی

پروژه درس هوش مصنوعی و سیستم‌های خبره

پیشبینی فروش—Deep Learning

استاد درس

دکتر غضنفری

دستیار استاد

دکتر حضرتی

دانشجو

سیدمهدی ولی‌زاده

۱ - چکیده	۱
۲ - مقدمه و بیان مسئله	۱
۳ - روش حل	۱
۳.۱ پیش پردازش داده‌ها و استخراج ویژگی‌ها	۱
۳.۲ طراحی شبکه عصبی و آموزش آن	۲
۴ - نتایج	۳

۱ - چکیده

مسئله پیش رو در این پژوهش پیشبینی فروش ماه آینده یک شرکت بر اساس فروش ۱۱۱۱ روز قبل آن بود که با استفاده از یادگیری عمیق و الگوریتم LSTM یک مدل یادگیری عمیق توسعه داده شد. ضمناً مدل فوق با روش K-Fold cross validation یا مقدار $K = 3$ آموزش داده شد و میانگین مربع خطای بدست آمده برابر با حدود ۰.۰۱ بیانگر آموزش خوب مدل فوق می‌باشد.

۲ - مقدمه و بیان مسئله

مسئله این پژوهش پیشبینی فروش یک شرکت با استفاده از داده‌های سری زمانی است. همانطور که در صورت مسئله گفته شد، متغیرهای مستقل برای پیشبینی را روز ماه، روز هفته، ماه سال، هفته سال و روز سال در نظر گرفتیم. برای حل مسئله طبق آنچه در صورت مسئله گفته شده، از شبکه عصبی و یادگیری عمیق به همراه لایه LSTM استفاده کردیم که به تفصیل مراحل آن در ادامه شرح داده خواهد شد.

۳ - روش حل

در این بخش از گزارش به تفصیل مراحل حل مسئله توضیح داده خواهد شد.

۳/۱ پیش پردازش داده‌ها و استخراج ویژگی‌ها

با توجه به این که داده‌ها را با فرمت مناسب در اختیار داشتیم، تنها لازم به استخراج ویژگی‌ها یا همان متغیرهای مستقل مسئله بود. برای راحتی کار ابتدا تمام تاریخ‌ها شمسی را به میلادی تبدیل کردیم تا بتوان به راحتی ویژگی‌های خواسته شده را با توابع کتابخانه‌های پایتون استخراج نمود. کد مربوط به این بخش را در اینجا می‌توانید در بلاک‌های ۴م تا ۹م فایل notebook مشاهده نمایید.

سپس به منظور این که داده‌ها باید وارد شبکه عصبی شوند، آن‌ها را با استفاده از یک scaler به بازه صفر تا یک مقیاس کردیم. لازم به ذکر است که به دلیل اختلاف بازه بسیار متغیر وابسته و متغیرهای مستقل ترجیح داده شد برای هر کدام از یک scaler مجزا استفاده شود. کد این بخش را می‌توانید در بلاک‌های ۱۲م تا ۱۵م مشاهده نمایید.

در انتهای این بخش لازم بود به منظور ورود داده‌ها به لایه LSTM که لایه ابتدایی شبکه ماست داده‌ها به اصطلاح reshape شوند. این به معنا است که یک پنجره زمانی با اندازه مشخص روی مجموعه داده ما حرکت می‌کند و هر بار یک پنجره زمانی خاص به لایه ورودی وارد می‌شود. اندازه پنجره زمانی با توجه به میزان داده‌های مسئله، مقدار روزهایی که باید پیشبینی شوند و ابعاد شبکه (چون انتخاب تعداد نوروها

لایه LSTM به این موضوع وابسته است) برابر با ۳۰ انتخاب شد. کدهای مربوط به این بخش نیز در بلاک‌های ۱۶ و ۱۷ قابل مشاهده است.

۳/۲ طراحی شبکه عصبی و آموزش آن

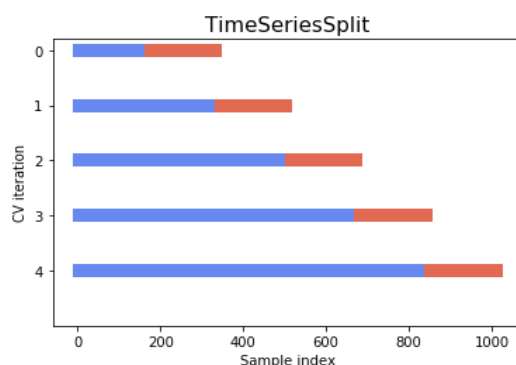
در شکل زیر خلاصه شبکه عصبی که خروجی کتابخانه keras پایتون است را مشاهده می‌کنید.

Layer (type)	Output Shape	Param #
lstm_3 (LSTM)	(None, 30)	4440
dropout_3 (Dropout)	(None, 30)	0
dense_3 (Dense)	(None, 1)	31
Total params: 4,471		
Trainable params: 4,471		
Non-trainable params: 0		

شکل ۱، خلاصه مدل آموزش داده شده

همانطور که در شکل نیز مشخص است لایه ابتدایی شبکه یک لایه LSTM با مقدار ۳۰ نورون است به منظور جلوگیری از خطر بیش برآزش یک dropout با مقدار ۰.۲ نیز قرار دادیم. با این کار نورون‌ها با احتمال ۰.۲ در هر تکرار ممکن است در نظر گرفته نشوند. این کار باعث کمتر شدن پیچیدگی شبکه شده و از بیش برآزش جلوگیری می‌کند. در نهایت نیز به لایه معمولی با مقدار ۱ نورون به منظور گرفتن خروجی قرار دادیم. لازم به ذکر است که تعداد نورون‌ها LSTM با در نظر گرفتن lag که در بخش قبل توضیح داده شد، انتخاب شده است.

چون صورت مسئله ما را به استفاده از رویکر cross validation ملزم کرده بود، داده‌ها نیز با همین منطق وارد شبکه شدند. در مسئله سری زمانی استفاده از cross validation مانند شکل زیر است.



شکل ۲، cross validation در مسئله سری زمانی

همانطور که شکل گویا است، در هر تکرار داده‌های ابتدایی به منظور آموزش و یک قسمت پس از آن به منظور آزمایش استفاده می‌شود. ما به دلیل تعداد recordهای کم مسئله تعداد Fold ها را برابر ۳ قرار دادیم. همچنین میزان داده‌های آزمایش در هر بخش برابر ۱۰۰ تعیین شد. (چون تعداد داده‌ها در هر بخش بسیار کم است)

همچنین به منظور جلوگیری از بیش برآزش از روش early stopping استفاده کردیم. این روش به این صورت است که در هنگام آموزش شبکه بخشی از داده‌ها را برای validation در نظر می‌گیرد و هنگامی خطا بر روی validation (بر خلاف داده‌های آموزش که همیشه رو به کاهش است) رو به افزایش می‌آورد، آموزش شبکه را متوقف می‌کند. ما تعداد epoch ها را به صورت پیشفرض برابر ۵۰۰ تعیین کردیم و با استفاده از early stopping گفتیم اگر تا ۱۰ epoch مقدار خطا روی validation بهتر نشد آموزش شبکه را موقوف کرده و بهترین مقدار پارامترها را برای شبکه به عنوان خروجی فرآیند آموزش در نظر بگیرد. لازم به ذکر است که تعداد داده‌های validation نیز برابر ۱۰۰ در نظر گرفته شد.

ضمناً به منظور آموزش بهتر شبکه داده‌هایی که برای آموزش شبکه استفاده می‌شدند (نه validation) به صورت بهم ریخته وارد شبکه کردیم. با استفاده از پارامتر shuffle. همچنین میزان batch size را با توجه به میزان اندک داده‌ها برابر ۱۶ قرار دادیم.

برای بهینه سازی شبکه از الگوریتم بهینه سازی Adam که در ادبیات متداول است با نرخ یادگیری ثابت ۰.۰۰۱ در نظر گرفته و تابع loss را برابر میانگین مربع خطا که یکی از توابع خطای رایج در مسئله رگرسیون است قرار دادیم.

کد مربوط به توضیحات فوق در بلاک‌های ۱۸ تا ۲۰ قابل مشاهده است.

۴ - نتایج

با توجه به این که تعداد Fold ها در روش cross validation برابر ۳ قرار داده شد، میزان خطا در هر کدام از این ۳ بخش به صورت زیر قابل مشاهده است.

جدول ۱، نتایج آموزش مدل

بخش	مقدار میانگین مربع خطا
بخش اول	۰.۰۱۵۲
بخش دوم	۰.۰۰۷۶
بخش سوم	۰.۰۱۱۲

جدول فوق گویای این است که مدل به خوبی آموزش دیده و میزان خطا اندک است.

به منظور پیش‌بینی ماه آینده نیز با استفاده از پایتون ۳۰ روز پس از آخرین روز داده شده در مجموعه داده‌ها ساخته شد؛ سپس ویژگی‌ها منطبق با مجموعه داده آموزش استخراج شده و داده‌ها در بازه صفر و یک مقیاس شد. سپس مطابق آن‌چه در بخش ۳-۱ توضیح داده شد داده‌ها آماده ورود به شبکه عصبی شده و خروجی مدل نیز گرفته شد. سپس خروجی با استفاده از scaler تعریف شده از قبل به مقیاس گذشته برگردانده شد که می‌توان آن را در جدول زیر مشاهده نمود.

میزان فروش	شماره روز
1809.2671	1
1839.1765	2
1818.8552	3
1760.4587	4
1669.5164	5
1545.556	6
1633.1686	7
1697.9066	8
1733.7761	9
1719.2046	10
1648.9055	11
1540.1945	12
1400.8114	13
1494.7985	14
1571.3604	15
1620.5121	16
1611.1039	17
1538.0114	18
1414.6427	19
1272.272	20
1374.2333	21
1502.8148	22
1610.588	23
1683.0647	24
1703.4943	25
1682.0061	26
1627.0371	27
1738.5524	28
1817.0245	29
1864.4706	30

کد مربوط به مطالب فوق نیز در بلاک‌های ۳۰ تا انتها قابل مشاهده است.