

## تمرین چهارم

**هدف:** آشنایی با شبکه‌های بازگشتی

**کد:** کد این فعالیت را می‌توانید به زبان پایتون و با استفاده از چارچوب‌های کاری تنسورفلو<sup>۲</sup> یا تورچ بنویسید.

**گزارش:** ملاک اصلی انجام فعالیت گزارش آن است و ارسال کد بدون گزارش فاقد ارزش است. برای این فعالیت یک فایل گزارش در قالب pdf تهیه کنید و در آن برای هر سوال، تصاویر ورودی، تصاویر خروجی و توضیحات مربوط به آن را ذکر کنید. سعی کنید توضیحات کامل و جامعی تهیه کنید.

**تذکر ۱:** مطابق قوانین دانشگاه هر نوع کپی برداری و اشتراک کار دانشجویان غیر مجاز بوده و شدیداً برخورد خواهد شد. استفاده از کدها و توضیحات اینترنت به منظور یادگیری بلامانع است، اما کپی کردن غیرمجاز است.

**تذکر ۲:** در تمام موارد، مجموعه داده‌های ورودی را به سه قسمت آموزش (۷۰ درصد)، اعتبارسنجی<sup>۱</sup> (۲۰ درصد) و آزمون (۱۰ درصد) تقسیم نمایید. آموزش شبکه را روی مجموعه داده آموزشی انجام دهید و تعداد تکرارها را با استفاده از مجموعه داده اعتبارسنجی بیابید. در نهایت برای ارزیابی عملکرد مدل از مجموعه داده آزمون استفاده نمایید.

**تذکر ۳:** برای ارزیابی عملکرد مدل‌ها در بخش دسته‌بندی از معیارهای دقت<sup>۲</sup> و ماتریس درهم‌ریختگی<sup>۳</sup> و در بخش رگرسیون از معیار میانگین مربعات خطا<sup>۴</sup> استفاده نمایید.

**تذکر ۴:** به منظور افزایش سرعت آموزش شبکه، می‌توانید از پلتفرم کولب گوگل<sup>۵</sup> استفاده نمایید. برای آشنایی با این پلتفرم به ویدیوهای آموزشی تدریس‌ار مراجعه فرمایید.

**راهنمایی:** در صورت نیاز می‌توانید سوالات خود را در خصوص پروژه از تدریس یار درس، از طریق ایمیل زیر بپرسید.

E-mail: ann.ceit.aut@gmail.com

**ارسال:** فایل‌های کد و گزارش خود را در قالب یک فایل فشرده با فرمت StudentID\_HW۰۴.zip تا تاریخ ۱۳۹۹/۱۰/۱۲ ارسال نمایید. شایان ذکر است هر روز تاخیر باعث کسر ۱۰٪ نمره خواهد شد.

در این تمرین قصد داریم عملکرد شبکه عصبی بازگشتی را در مدل‌سازی دنباله‌های زمانی مورد بررسی قرار دهیم. مجموعه داده مورد استفاده در اینجا، مقدار شاخص صنعت خودروسازی در بازار بورس ایران در سال‌های اخیر است که به فایل تعریف پروژه پیوست شده است. هر سطر از فایل اکسل مجموعه داده، اطلاعات شاخص در یک روز را نمایش می‌دهد. در هر سطر می‌توانید، قیمت ابتدای

<sup>۱</sup> Validation

<sup>۲</sup> Accuracy

<sup>۳</sup> Confusion matrix

<sup>۴</sup> Mean squared error

<sup>۵</sup> Google Co-lab: <https://colab.research.google.com/>

روز، بیشترین قیمت در طول روز، کمترین قیمت در طول روز، قیمت پایانی (میانگین قیمت معامله شده در طول روز) و آخرین قیمت را مشاهده نمایید. همین‌طور اطلاعاتی از جمله حجم معاملات (تعداد سهام معامله شده) و ارزش کل معاملات در طول روز نیز در هر سطر از این مجموعه داده قرار داده شده است.

برای انجام این تمرین مراحل زیر را اجرا کرده و نتایج را در فایل گزارش ذکر نمایید.

۱- قیمت پایانی شاخص خودروسازی را در یک نمودار رسم نمایید. در این نمودار روزهایی را که قیمت پایانی روز بیشتر از قیمت پایانی روز قبل است را با رنگ سبز و روزهایی که قیمت پایانی کمتر از قیمت پایانی در روز قبل است را با رنگ قرمز نمایش دهید. این نمودار را به سه قسمت آموزش، اعتبارسنجی و آزمون تقسیم نمایید. توجه نمایید در تقسیم داده‌ها، نباید تصادفی عمل شود. داده‌های سه ماه آخر را به عنوان آزمون، سه ماه قبل از آن را به عنوان داده‌های اعتبارسنجی و مابقی را به عنوان داده‌های آموزشی استفاده نمایید.

۲- سوال ۱ را یک بار دیگر برای پنجره زمانی یک ماهه حل نمایید. در این حالت، قیمت پایانی شاخص در هر ۲۰ روز را در یک گام زمانی رسم نمایید. در مواقعی که قیمت پایانی یک ۲۰ روز از قیمت پایانی ۲۰ روز قبل بیشتر باشد، نمودار را با رنگ سبز و در غیر این صورت با رنگ قرمز نمایش دهید. این نمودار را به سه قسمت آموزش، اعتبارسنجی و آزمون تقسیم نمایید.

توجه نمایید در این حالت داده‌های شش ماه آخر را به عنوان آزمون، سه ماه قبل را به عنوان اعتبارسنجی و مابقی را به عنوان آموزشی در نظر بگیرید.

۳- در این قسمت قصد داریم با استفاده از یک شبکه عصبی، صعودی یا نزولی بودن شاخص در گام زمانی بعدی را پیش‌بینی نماییم. برای این کار، یک شبکه عصبی LSTM را استفاده نمایید. این شبکه را یک بار روی شاخص روزانه، داده‌های سوال ۱، و یک بار روی شاخص ماهانه، داده‌های سوال ۲، آموزش داده و نتایج را با هم مقایسه نمایید. توجه نمایید در این قسمت فقط باید از قیمت‌های پایانی روزهای قبل استفاده نمایید. پارامترهای بهینه برای شبکه را یافته و نتایج آزمایشات را در گزارش ذکر نمایید. عملکرد شبکه در پنجره زمانی روزانه بهتر است یا ماهانه؟ فکر می‌کنید این تفاوت در عملکرد از کجا ناشی می‌شود؟ در هر یک از حالات، مجموعه داده آزمون و آموزشی را به طور جداگانه رسم نموده و گام‌های زمانی که شبکه شما به عنوان گام زمانی صعودی پیش‌بینی کرده است را با رنگ سبز و مابقی گام‌ها را با رنگ قرمز نمایش دهید.

۴- در این سوال، علاوه بر قیمت‌های پایانی در روزهای قبل، قیمت‌های آغازین، کمینه و بیشینه قیمت، حجم و ارزش معاملات را نیز در نظر بگیرید. شبکه را یک بار دیگر روی هر دو مجموعه داده به طور مجزا آموزش داده و عملکرد شبکه‌ها را با یکدیگر مقایسه نمایید. همین‌طور عملکرد بهترین مدل در این سوال را با عملکرد بهترین مدل در سوال ۳ نیز مقایسه نمایید. چه نتیجه‌ای از این مقایسه‌ها می‌گیرید؟ نتایج تمام آزمایشات را به همراه تحلیل‌های خود در فایل گزارش ذکر نمایید.

توجه نمایید برای محاسبه بیشینه و کمینه قیمت در یک ماه، کافیست به ترتیب بیشترین مقدار بیشینه قیمت و کمترین مقدار کمینه قیمت در روزهای داخل ماه را در نظر بگیرید. همین‌طور آخرین قیمت پایانی در طول یک ماه را به عنوان قیمت پایانی آن ماه در نظر بگیرید. حجم و ارزش معاملات در یک ماه نیز معادل مجموع حجم و ارزش معاملات در روزهای هر ماه است.

۵- شبکه سوال ۳ را به نحوی تغییر دهید که به جای پیش‌بینی صعودی یا نزولی بودن فقط در یک روز آینده، تمام ۲۰ روز آینده را پیش‌بینی نماید. ساختار شبکه جدید را به همراه ارزیابی عملکرد آن و نتایج آزمایشات در فایل گزارش ذکر نمایید. تفاوت در عملکرد این مدل و مدل سوال ۳ در چیست؟ فکر می‌کنید این مشکل را چگونه می‌توان حل کرد؟

