

## « تمرین درسی »

درس	عنوان	مهلت	نمره	نوع
مباحث ویژه (یادگیری ماشین)	تمرین ۳	۸ آذر ۱۴۰۱ - ۲۳:۵۹	۲/۵	الزامی

- I. **هدف آموزشی تمرین:** در این تمرین قصد داریم با استفاده از یک مجموعه داده با نحوه عملکرد یادگیری ماشین نظارتی بیشتر آشنا شویم. برای پیش برد این یادگیری از الگوریتم درخت تصمیم برای تجزیه و دسته بندی مجموعه داده کمک خواهیم گرفت.
- II. **نحوه ارسال تمرین:** پیاده سازی انجام شده را در قالب یک فایل Jupyter notebook یا پایتون (.py) به همراه یک گزارش کار در قالب Word و PDF. محتوای نهایی تمرین به شکل فشرده در بخش تکالیف درسی در سامانه سمیاد ارسال شود. تحويل گزارش کار الزامی است.
- III. **شرح و محتوای تمرین:**

درخت تصمیم یک نمایش ساده برای طبقه بندی داده ها است و یک روش یادگیری ماشینی تحت نظارت است که در آن داده ها به طور مداوم بر اساس یک پارامتر خاص تقسیم می شوند. تجزیه و تحلیل درخت تصمیم می تواند به حل مشکلات طبقه بندی و رگرسیون کمک کند. الگوریتم درخت تصمیم یک مجموعه داده را به زیر مجموعه های کوچک تر تقسیم می کند. در حالی که در همان زمان، درخت تصمیم مرتبط به صورت تدریجی توسعه می یابد. درخت تصمیم شامل گره ها (که مقدار یک ویژگی خاص را آزمایش می کند)، یال ها/شاخه ها (که با نتیجه آزمایش مطابقت دارند و به گره یا برگ بعدی متصل می شوند) و گره های برگ (گره های پایانی که نتیجه را پیش بینی می کنند) تشکیل شده است که آن را به یک ساختار کامل تبدیل می کند. در این تمرین، از بسته یادگیری scikit برای پیاده سازی درخت تصمیم در پایتون استفاده کنید.

برای تجزیه و تحلیل، از مجموعه داده ضمیمه شده که در زمینه علوم پزشکی است استفاده نمایید. بر اساس متغیرهای ثبت شده در مجموعه داده ها، هدف پیش بینی آن است که آیا بیمار مبتلا به دیابت است یا خیر؟ این اطلاعات شامل متغیرهای پیش بینی کننده مانند BMI بیمار، جزئیات بارداری، سطح انسولین، سن و غیره است.

- موفق باشید...