

## تمرین سوم

---

**هدف:** آشنایی با شبکه‌های توابع پایه شعاعی (تپش)

**کد:** کد این فعالیت را به زبان پایتون 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:** مجموعه‌های داده مورد استفاده را به جز در مواردی که صریحاً در صورت سوال ذکر شده باشد، حتماً قبل از استفاده بصورت تصادفی به سه بخش آموزش (۷۰ درصد داده‌ها)، آزمون (۲۰ درصد داده‌ها) و اعتبارسنجی (۱۰ درصد داده‌ها) تقسیم نمایید.

**تذکر ۳:** مدل‌های تخمین گر را بر اساس معیار میانگین مربعات خطا ارزیابی نمایید.

**راهنمایی:** در صورت نیاز میتوانید سوالات خود را در خصوص پروژه از تدریس‌یار 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\_HW۰۳.zip تا تاریخ ۱۴۰۰/۰۹/۱۲ ارسال نمایید. شایان ذکر است هر روز تاخیر باعث کسر ۱۰٪ نمره خواهد شد.

---

در این تمرین قصد داریم با استفاده از شبکه‌های تپش، میزان مصرف سوخت<sup>۱</sup> اتوموبیل‌های مختلف را بر اساس ویژگی‌های موجود تخمین بزنیم. مجموعه داده مورد استفاده در این تمرین، مجموعه داده Auto MPG است که شامل ۳۹۸ رکورد از اتوموبیل‌های مختلف است. داده‌های موجود و اطلاعات بیشتر در خصوص این مجموعه داده را می‌توانید از طریق لینک زیر مشاهده نمایید.

<https://archive.ics.uci.edu/ml/datasets/Auto+MPG>

برای انجام این تمرین به سوالات زیر پاسخ دهید.

۱- مجموعه داده مورد استفاده را بارگذاری نموده و داده‌های نامعلوم<sup>۲</sup> را با روش مناسب جایگزین نمایید. نحوه جایگزینی این

داده‌ها را در فایل گزارش بطور کامل توضیح دهید.

۲- میزان مصرف سوخت را به ازای هر یک از ویژگی‌های موجود در نمودار جداگانه‌ای رسم نمایید. نمودار داده‌ها را در فایل گزارش رسم نمایید.

---

<sup>۱</sup> Miles per gallon (MPG)

<sup>۲</sup> Missing values

- ۳- یک شبکه عصبی تپش را یک بار با استفاده از الگوریتم  $OLS^3$  و یک بار با استفاده از خوشه‌بندی توسط الگوریتم Kmeans آموزش دهید. عملکرد شبکه‌ها در این دو حالت را با هم مقایسه نموده و بگویید به نظر شما کدام یک از این الگوریتم‌ها برای حل مساله مورد نظر مناسب‌تر است و چرا؟ آزمایش زیر را روی هر یک از شبکه‌ها انجام داده و ضمن توضیح نتایج آزمایش در فایل گزارش، بر اساس نتایج آن به سوال پاسخ دهید.
- عملکرد شبکه را با تعداد نورون‌های لایه مخفی ۵، ۱۰، ۵۰، ۱۰۰، ۱۵۰ بررسی نمایید. افزایش تعداد نورون‌ها چه تاثیری بر عملکرد مدل دارد؟ پاسخ خود را با دلایل روشن تحلیل کنید.
- ۴- یک شبکه عصبی پرسپترون چندلایه را برای رگرسیون مجموعه داده ضمیمه شده آموزش دهید. ساختار این شبکه را در فایل گزارش، توضیح دهید. آزمایشات زیر را بر روی این شبکه انجام داده و نتایج آن‌ها را در فایل گزارش تحلیل نمایید.
- عملکرد شبکه را با تعداد لایه‌های ۱، ۲، ۳، ۵، ۱۰، ۵۰، ۱۰۰ بررسی نمایید. افزایش تعداد لایه‌های چه تاثیری بر عملکرد مدل دارد؟ پاسخ خود را با دلایل روشن تحلیل کنید.
  - عملکرد شبکه را با تعداد نورون‌های ۵، ۱۰، ۵۰، ۱۰۰، ۱۵۰ با در نظر گرفتن بهترین تعداد لایه‌ها که از آزمایش قبل حاصل شده، بررسی نمایید. افزایش تعداد نورون‌ها چه تاثیری بر عملکرد مدل دارد؟ پاسخ خود را با دلایل روشن تحلیل کنید.
  - عملکرد بهترین مدل را با نرخ‌های یادگیری ۰.۰۰۰۱، ۰.۰۰۰۱، ۰.۰۰۱، ۰.۰۱، ۰.۰۹ و ۱ بررسی نمایید. افزایش نرخ یادگیری چه تاثیری بر عملکرد مدل دارد؟ پاسخ خود را با دلایل روشن تحلیل کنید.
- ۵- براساس نتایج بدست آمده، عملکرد شبکه تپش را با شبکه پرسپترون چندلایه مقایسه نموده و نقاط قوت و ضعف هر یک از این مدل‌ها در رگرسیون را بطور کامل توضیح دهید.

موفق باشید

---

<sup>۳</sup> Orthogonal Least Square