

به نام خالق عقل و خرد



جلسه دوم

یادگیری ماشین

نیم سال دوم ۱۴۰۴ - ۱۴۰۳

دانشکده ریاضی و آمار

دانشگاه اصفهان

اسفند ماه ۱۴۰۳

استاد درس:

دکتر فاطمه منصوری

عنوان تمرین : پیاده‌سازی رگرسیون خطی و رگرسیون چند جمله‌ای

دستورالعمل :

۱. یک دیتاست ساختگی ایجاد کنید. برای این کار یک تابع بنویسید که (عنوان تابع در فایل ژوپیتتر قرار داده شده است) دیتاست را مطابق زیر ایجاد کند. دیتاست شامل ۱۰۰۰ نمونه باشد. دیتاست شامل یک ویژگی مستقل x است. مقدار x بین ۳- تا ۳ تغییر می‌کند و ارتباط بین متغیر وابسته و متغیر ساختگی مطابق زیر باشد.

$$y = 0.2x^3 - 0.5x^2 + 3x + 10 + noise$$

۲. با استفاده از یک scatter plot ارتباط بین x و y را مشاهده نمایید.

۳. داده‌ها را به train و test تقسیم کنید. علاوه بر آن یک مجموعه تست در بازه ۵ تا ۷ به نام out_of_range_test ایجاد کنید.

۴. برای ویژگی polynomial feature ایجاد کنید و یک مدل polynomial regression به داده‌ها اختصاص دهید.

۵. تابع train_models را مطابق آنچه در فایل ژوپیتتر خواسته شده است پیاده‌سازی کنید و سپس با استفاده از آن مدل‌ها با درجات مختلف ایجاد کرده و نمایش دهید.

۶. تابع evaluate_models نوشته شده در فایل ژوپیتتر را کامل کنید به صورتیکه به ازای مدل‌های ساخته شده با درجات مختلف مقدار خطا را محاسبه کرده و در آخر لیستی از خطاها برگرداند.

۷. خطا به ازای درجات مختلف برای مجموعه داده آموزش، مجموعه داده تست و مجموعه داده تست خارج از بازه را به دست آورده و این خطاها را به ازای مقادیر مختلف درجات روی یک نمودار رسم کنید.

۸. نمودارهای به دست آمده از مرحله قبل را تحلیل کنید.