

سؤال اول پیاده سازی:

مقدمه: یک مسئله ی رگرسیون خطی است که هدف آن پیش بینی قیمت خانه بر اساس چندین ویژگی است. برای حل این مسئله، مراحل زیر را پیاده سازی کرده ایم:

حل مسأله و پیاده سازی آن:

ابتدا، داده های آگهی فروش خانه را از یک فایل CSV به محیط برنامه نویسی پایتون وارد کرده و با استفاده از کتابخانه پانداس، یک داده ساختاری (dataframe) از آن ساخته ایم.

سپس، داده ها را پیش پردازش کرده شامل: (نمایش ۵ سطر اول داده ها/نمودار هیستوگرام/ نمایش اطلاعات داده ها/ بدست آوردن مقدار یکتا برای ستون آدرس/حذف سطرهایی که داده های پرت دارند/ تبدیل کردن مقدار غیر عددی یک ستون به مقدار عددی و...)

و سپس بعد از پیش پردازش داده ها را به دو بخش آموزش و تست تقسیم کرده ایم. بخش آموزش شامل 80 درصد داده ها و بخش تست شامل 20 درصد داده ها است. ما از کتابخانه sklearn برای تقسیم داده ها استفاده کرده ایم.

در نهایت، یک مدل رگرسیون خطی با استفاده از کتابخانه sklearn ساخته و با داده های آموزش، آموزش داده ایم.

نتیجه:

سپس، با استفاده از داده های تست، قیمت خانه را پیش بینی کرده و مقدار خطای مربعات میانگین (MSE) و مقدار خطای قدر مطلق میانگین (MAE) را به عنوان معیار ارزیابی مدل محاسبه کرده ایم.