

هدف این تمرین آشنایی شما با مفاهیم اصلی دسته‌بندی است. شما باید ضمن اجرا گرفتن از کد دسته‌بندی استفاده شده در کلاس، یک بار کد را با دسته‌بند KNeighborsClassifier اجرا کنید.

در این تکلیف باید به موارد زیر پاسخ دهید.

۱- بالاترین میزان صحت (accuracy) که از این دو دسته‌بند حاصل می‌شود چه میزان است. روش ارزیابی 10-fold cross validation است.

۲- معیارهای precision و recall و F1 و accuracy را برای نتایج نهایی مورد ۱ محاسبه کنید. به نظر شما کدام یک از این ۴ معیار برای این مسئله اهمیت بیشتری دارد؟ حال تلاش کنید بالاترین نتایج را در معیار انتخابی (با تغییر پارامترها) بدست آورید.

نکات:

۱- در خصوص پیاده‌سازی و پارامترهای Naive bayes می‌توانید با یک جستجوی ساده مطالب مفیدی پیدا کنید.

۲- خروجی کار را در قالب یک فایل Jupiter تحویل دهید که جواب سوالات کدها و نتایج اجرا در آن آمده است.

موفق باشید.