

تمرین سری اول درس هوش محاسباتی

مهلت ارسال تمارین: روز سه شنبه 21 فروردین 1398

نحوه ارسال : تمامی فایل های خود را در فایل .zip با شماره دانشجویی خود قرار داده و به پست الکترونیکی زیر ارسال کنید.

ci.class.12+seri1@gmail.com

هر فایل باید شامل یک بخش کد و یک فایل PDF. گزارش تمرین باشد. لطفا گزارش به فرمت مقاله نوشته شود (فرمت مقاله ضمیمه شده است)

الف) هر یک از 2 طبقه بند، 1NN و BAYES را بر روی مجموعه داده ی معرفی شده پیاده سازی کنید و در گزارش خود نتایج اجرای آنها بر روی هر کدام از مجموعه داده ها را مقایسه کرده و دلیل عملکرد بهتر یا بدتر آنها را توضیح دهید.

ب) برای 2 طبقه بند بالا به ازای پارامترهای مختلف هر الگوریتم نیز نتایج را تحلیل نمایید.

مانند : شعاع همسایگی یا تعداد آن). در نهایت شرح دهید که هر کدام از این الگوریتم ها روی کدام مجموعه داده بهتر کار می کنند و به چه دلیل ؟

مجموعه داده های معرفی شده :

(۱) یک مجموعه داده ی دو بعدی دو کلاس با توزیع نرمال ایجاد کنید. انتخاب تعداد نمونه های آزمایش

و ارزیابی بر عهده ی خود دانشجویان است. پارامترهای این توزیع ها برای کلاس اول و دوم به ترتیب عبارتست از :

کلاس اول - میانگین : $[0, 0]$ ، ماتریس کواریانس $[0.5, 0.3, 0.3, 1]$

کلاس دوم - میانگین : $[1, 2]$ ، ماتریس کواریانس $[0.25, 0.3, 0.3, 1]$

(۲) یک مجموعه داده ی دو بعدی سه کلاس همانند بالا با توزیع نرمال ایجاد کنید. در این مورد نیز

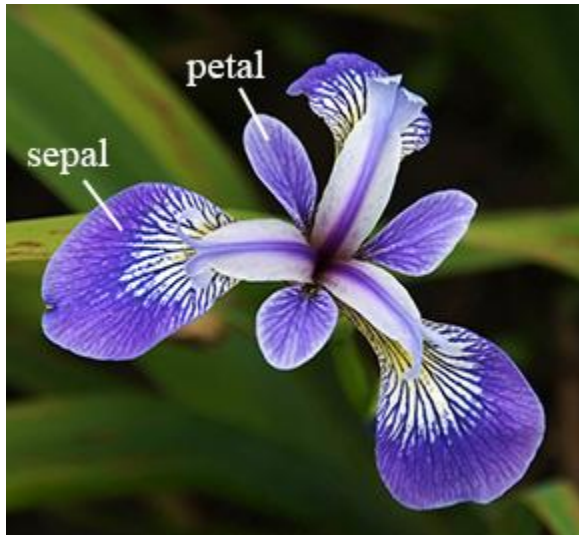
انتخاب تعداد نمونه های آزمایش و ارزیابی بر عهده ی خود دانشجویان است. پارامترهای توزیع ها به قرار زیر است:

کلاس اول - میانگین : $[-1.5, 1]$ ، ماتریس کواریانس $[0.5, 0.3, 0.3, 1.1]$

کلاس دوم - میانگین : $[0.5, 0]$ ، ماتریس کواریانس $[0.25, 0.3, 0.3, 1]$

کلاس سوم - میانگین : $[-1.5, -2.25]$ ، ماتریس کواریانس $[0.25, 0.3, 0.3, 0.7]$

۳) مجموعه داده iris که مجموعه داده گل زنبق است که براساس petal و sepal به گونه های متفاوت تقسیم می شوند.



این مجموعه داده را می توان از لینک زیر دانلود کنید:

<https://archive.ics.uci.edu/ml/datasets/Iris>

راهنمایی: کسانی که از زبان پایتون استفاده می کنند میتوانند از لینک زیر هم برای دسترسی به داده استفاده کنند.

https://scikit-learn.org/stable/modules/generated/sklearn.datasets.load_iris.html

سوال امتیازی: مجموعه داده اعداد فارسی دست نویس جمع آوری شده که دارای 10 کلاس از اعداد از صفر تا 10 است. این دو روش را بر روی این مجموعه اعمال کرده و نتیجه را تحلیل کنید. (این مجموعه داده ضمیمه شده است)