

مقدمه ای بر یادگیری ماشین

نیمسال اول ۱۳۹۹-۱۴۰۰

مدرس: صابر صالح

تمرین سری اول

● مهلت تحویل تمرین های عملی: ۱۳۹۹/۰۸/۰۱ ●

● تمرین عملی

توضیحات

به پیوست تمرین تعدادی فایل ipynb آمده است که هر کدام مربوط به یک بخش از تمرین می باشد. شما می بایست قسمت های مختلف تمرین را در این فایل ها تکمیل کرده و با همین فرمت ارائه شده تحویل دهید. توجه داشته باشید که هر بخش از تمرین که با فرمت ارائه شده متفاوت باشد تصحیح نخواهد شد. در برخی قسمت های تمرین ممکن است که نیاز باشد توضیحات مختصری ارائه کنید. در این موارد، توضیحات را به صورت Markdown در نوت بوک هر بخش ذکر کنید. در نهایت همه فایل های تمرین را به صورت یک فایل فشرده به فرمت ML-HW1-9xxxxxxx.zip تحویل دهید.

آشنایی با numpy

یکی از کتابخانه های پر کاربرد پایتون در یادگیری ماشین و علوم داده کتابخانه numpy می باشد. در این تمرین، به منظور آشنایی با این کتابخانه تعدادی عملیات ماتریسی و ریاضیاتی بر روی آرایه ها انجام می دهیم. توجه داشته باشید در بخش های مختلف تمرین ورودی هر بخش خروجی عملیات بخش های قبلی بوده و ترتیب اجرای عملیات های مختلف در جواب نهایی تاثیرگذار می باشد. فایل مربوط به این قسمت از تمرین، فایل Problem1.ipynb می باشد.

آشنایی با pandas

یکی دیگر از کتابخانه های پر کاربرد پایتون در علوم داده کتابخانه pandas می باشد. در این بخش به منظور آشنایی بیشتر با این کتابخانه و قابلیت های آن از دیتاست IMDB که به پیوست تمرین آمده است استفاده می کنیم. ابتدا این دیتاست را لود کرده و سپس بخش های مختلف خواسته شده را انجام دهید. فایل مربوط به این قسمت از تمرین، فایل Problem2.ipynb می باشد.

بصری سازی^۱ در پایتون

یکی از مهارت های مهم در انجام پروژه های یادگیری ماشین توانایی بصری سازی داده می باشد. در این بخش با برخی از قابلیت ها و ابزارهای بصری سازی در پایتون آشنا می شویم. بدین منظور در این بخش از دیتاست iris استفاده می کنیم. ابتدا این دیتاست را از کتابخانه sklearn لود کنید و سپس موارد خواسته شده را انجام دهید. فایل مربوط به این قسمت از تمرین، فایل Problem3.ipynb می باشد.

^۱Visualization