



دانشگاه صنعتی شریف

دانشکده مهندسی برق

آزمایشگاه برنامه نویسی پایتون

آزمایش ۶

در این آزمایش در ابتدا با برنامه نویسی شی گرا^۱ آشنا می شویم. سپس برخی از توابع پکیج‌های Pandas، matplotlib و numpy و scipy را با هم می‌بینیم.

۱ مقدمه ای بر برنامه نویسی شی گرا

۱-۱ کلاسی با نام Set برای تعریف مجموعه‌ها پیاده سازی کنید. هر نمونه از این کلاس، باید شامل تعدادی عضو باشد (برای سادگی فرض کنید اعضا می توانند از نوع عدد یا رشته باشند) و توابع و عملگرهای زیر روی آن تعریف شوند. دقت کنید برای نگهداری اعضای مجموعه باید از لیست استفاده کنید. عملکرد کلاسی که تعریف کرده‌اید را بررسی نمایید.

عملکرد	تابع یا عملگر
یک مجموعه ی تهی ایجاد می کند.	"a = Set()"
در صورتی که x عضو a نباشد، آن را به a اضافه می کند.	"a.add(x)"
در صورتی که x عضو a باشد، آن را از a حذف می کند.	"a.delete(x)"
یک مجموعه که به ترتیب اجتماع و اشتراک a و b است ایجاد می کند.	"a & b" و "a b"
اگر دو مجموعه یکسان باشند True و اگر یکسان نباشند False بر می گرداند.	"a == b"

۲ Pandas

Pandas کتابخانه‌ای متن باز برای پایتون است که ساختارهای داده‌ای با کارایی بالا برای تجزیه و تحلیل داده ها را ممکن می‌سازد. در این آزمایش تمرکز روی یادگیری Dataframe و استفاده آن برای خواندن و ایجاد تغییر در یک دیتاست در فایل CSV می‌باشد.

۱-۲ لیست indexes را از اعداد یک تا ۶ بسازید. متغیر columns را نیز به همین ترتیب از حروف A تا D بسازید. سپس متغیر df را که یک Dataframe با ۴ ستون و ۶ سطر است از با استفاده از اعداد تصادفی بسازید. در حین ساختن این دیتافریم columns را برابر متغیر columns و index را برابر متغیر indexes قرار دهید. حال با دستور print این متغیر را مشاهده نمایید.

۱-۲ با ویژگی‌های head و tail دوسطر اول و آخر دیتافریم را نمایش دهید.

۲-۲ دستورهای df.values و df.columns و df.index و df.describe را چاپ کنید و نتایج و کاربرد هر دستور را تحلیل کنید.

¹ Object Oriented Programming

۳-۲ با دستور مناسب این دیتا فریم را بر اساس ستون B و به ترتیب نزولی و در صورتی که مقادیر در این ستون مساوی بودند بر اساس صعودی و نسبت به ستون C مرتب کنید. به صورتی مرتب سازی را انجام دهید که همزمان در متغیر df ذخیره سازی هم صورت بگیرد.

۴-۲ متغیر S را با استفاده از pd.Series و از اعداد ۱ تا ۶ بسازید. ستون F را با متغیر S به df اضافه کنید.

۵-۲ با استفاده از df.at یا df.iat سطر سوم و پنجم ستون F را برابر np.nan قرار دهید. حال با استفاده از df.loc یا df.iloc سه سطر اول ستون های A و F را نمایش دهید.

۶-۲ متغیر df\ را برابر متغیر df بدون خانه های NaN قرار دهید. (برای این کار میتوانید از df.dropna استفاده کنید).

۷-۲ متغیر df۲ را برابر متغیر df با پر کردن خانه های NaN با میانگین ستون آخر قرار دهید. (برای این کار میتوانید از df.fillna استفاده کنید).

۳ matplotlib

پکیج matplotlib ابزاری بسیار قوی برای ترسیم انواع نمودارها و گراف ها در پایتون است. در این قسمت با استفاده از این پکیج توزیع داده های یک Dataframe را ترسیم می کنیم.

۱-۳ فایل "train.csv" را خوانده و در یک Dataframe ذخیره کنید. هیستوگرام داده ها را بر اساس ApplicantIncome و در ۵۰ بین (bin) رسم کنید. نمودار داده ها را این بار با نمودار boxplot در حالتی که ستون های آن بر اساس ApplicantIncome و محور افقی آن بر اساس Education می باشد رسم کنید.

۲-۳ با استفاده از plt.figure یک figure ایجاد کنید. حال یک subplot اضافه کرده و در آن تعداد Applicant ها را بر اساس Credit_History نمایش دهید. (محور X را Credit_History و محور Y را تعداد Applicant ها قرار دهید). حال یک subplot دیگر اضافه کنید و در آن احتمال دریافت وام (Loan_Status) را بر اساس Credit_History رسم کنید. (ابتدا باید Loan_Status را از حالت categorical به numerical تبدیل کنید).

۴ Numpy & Scipy

پکیج‌های `numpy`^۲ و `scipy`^۳ برای محاسبات عددی و کاربردهای ریاضیاتی استفاده می‌گردند. همچنین پکیج `matplotlib`^۴ ابزاری بسیار قوی برای ترسیم انواع نمودارها و گراف‌ها در پایتون است. با استفاده از پکیج‌های معرفی شده، عملیاتی که در ادامه می‌آید را انجام دهید و خروجی را مشاهده کنید.

۱-۴ آرایه‌های `numpy` نسبت به لیست‌های پایتون از نظر حافظه و سرعت عملکرد بهتری دارند و قابلیت‌های کاربردی بیشتری نیز در اختیار می‌گذارند. یک آرایه ۶ عضوی با کمک تابع `arange` ایجاد نمایید و با کمک تابع `reshape` آن را به صورت 3×2 در آورید. سپس یک آرایه‌ی دیگر با اندازه‌ی 3×3 ایجاد نمایید و با کمک `stack` دو آرایه را زیر هم قرار دهید تا به یک آرایه‌ی 3×5 تبدیل شوند. با استفاده از `ravel` این آرایه را به یک آرایه‌ی مسطح تبدیل کنید و با استفاده از `argmax` اندیس بزرگترین عضو آن را بیابید.

۲-۴ ماتریس‌هایی به صورت زیر تشکیل دهید و با کمک `linalg`^۵ عملیاتی که ذکر شده را روی آنها انجام دهید.

ماتریس	عملیات
$A = \text{np.mat}([[3, 1, 4], [1, 5, 9], [2, 6, 5]])$ $b = \text{np.mat}([2, 4, -1])$	$A^T, A^{-1}, A $ $\text{eig}(A), \text{svd}(A), \text{qr}(A)$ Solve: $Ax=bT$

۳-۴ ابتدا یک بردار تصادفی t با 1000 عضو ایجاد کنید و آن را به بازه $-10\pi \leq t \leq 10\pi$ تصویر کنید. سپس بردار ایجاد شده را مرتب نمایید و به ازای نقاط این بردار تابع $y = \text{sinc}(t)$ را محاسبه نمایید. در مرحله‌ی بعد تبدیل فوریه $z = \text{fft}(y)$ را محاسبه کنید. در نهایت با کمک `pyplot` گراف تابع اصلی و اندازه‌ی تبدیل فوریه آن را در یک شکل و در قالب دو `subplot` زیر هم رسم نمایید. برچسب گذاری نمودار را به نحو صحیح انجام دهید و گراف حاصل را به صورت فایل `'figure.png'` ذخیره نمایید.

^۲ <http://www.numpy.org>

^۳ <https://www.scipy.org>

^۴ <https://matplotlib.org>

^۵ <https://docs.scipy.org/doc/scipy/reference/linalg.html>