



دانشگاه تهران، دانشکده ریاضی، آمار و علوم کامپیوتر

درس: یادگیری ماشین آماری

استاد درس: دکتر امینی

تمرین مقدماتی درس یادگیری ماشین آماری

مهلت تحویل: پایان روز چهارشنبه ۲ آبان ۱۴۰۳

توجه:

- کد باید به صورت یک فایل notebook شامل نتایج اجرا و در زبان پایتون باشد.
- علاوه بر کد، گزارش کاری که شامل توضیحات در مورد کد می باشد نیز باید ضمیمه گردد.
- فایل گزارش کار باید به فرمت pdf و به زبان فارسی و یا انگلیسی باشد.
- نمره این تمرین از ۱۰ است.

(۱) تابعی به نام `sum_primes` بنویسید که دو عدد صحیح `start` و `end` را دریافت و مجموع اعداد اول در بازه `[start, end]` را برگرداند.

**راهنمایی:** به عنوان مثال، اگر این تابع به صورت `sum_primes(start = 1, end = 6)` فراخوانی گردد مقدار خروجی ۱۰ باید برگردانده شود. (۱ نمره)

(۲) فرض کنید دنباله ای صعودی از اعداد طبیعی بزرگتر از یک داریم که بعضی از مقادیر آن تکراری است. به عنوان مثال دنباله زیر را در نظر بگیرید:

`input = [1, 1, 1, 2, 2, 3, 3, 3, 3, 3, 4]` (۱)

برای این دنباله قصد داریم فقط اولین اعداد طبیعی منحصر به فرد را نگه داریم و مقادیر تکراری را با عدد صفر جایگزین کنیم. پس برای دنباله بالا، خروجی زیر را انتظار داریم

`output = [1, 0, 0, 2, 0, 3, 0, 0, 0, 0, 4]`

تابعی به نام `gen_mask` بنویسید که یک لیست از اعداد به شکل (۱) را بعنوان ورودی دریافت و خروجی به صورت (۲) را برگرداند. (حالتی که دنباله وردی صعودی نباشد را در نظر گرفته و در صورت صعودی نبودن پیام خطای مناسب را برگردانید). (۲ نمره)

(۳) یک رشته را آئینه ای گوئیم اگر از دو سمت راست و چپ به یک صورت خوانده شود. رشته `radar` و یا ۱۲۳۲۱ بعنوان مثال آئینه ای ولی رشته `book` آئینه ای نیست. تابعی به نام `is_palindrome` بنویسید که یک رشته دریافت و در صورت آئینه بودن رشته مقدار `True` و در غیر اینصورت مقدار `False` را برگرداند. (۱ نمره)

(۴) از تابع سوال ۳ استفاده کرده و مجموع اعداد آئینه ای در بازه `[1, 1_000_000]` را محاسبه کنید. (۱ نمره)

(۵) برنامه بنویسید که یک فایل متنی را دریافت کند و فراوانی کلمات را محاسبه و به صورت یک دیکشنری که در آن کلید ها کلمات و مقادیر فراوانی کلمات میباشد ذخیره کند. و در آخر ۵ کلمه پر تکرار را به تعداد دفعات رخداد در متن نمایش دهد. این برنامه را روی فایل sample.txt اجرا کنید. (۱ نمره)

(۶) تابعی بنویسید که عدد طبیعی M را دریافت و مقدار زیر را محاسبه و برگرداند.

$$f(M) = \left[ 1 + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \dots + \frac{1}{\sqrt{M}} \right]$$

در این تابع منظور از [] جز صحیح میباشد. مثلاً [2.6]=2 (۱ نمره)

(۷) در فایل balance.txt موجودی حساب اشخاص در چهار بانک ملت، تجارت، رفاه و ملی ثبت شده است. اطلاعات این فایل به این صورت ذخیره شده است که بعد از یک سطر استار در سطر بعدی به ترتیب با فاصله نام بانک، نام خانوادگی صاحب حساب و موجودی ذخیره شده است. برنامه بنویسید که این فایل را خوانده و اطلاعات حساب اشخاص را در چهار دیکشنری به نام های meli, refah, tejarat و meli ذخیره کند. (۱ نمره)

(۸) یک دیکشنری به نام all\_banks ایجاد کنید و اطلاعات هر چهار دیکشنری اشاره شده در سوال ۷ را ترکیب کنید. (۱ نمره)

**راهنمایی:** در صورتی که شخصی که فقط در یکی از این بانک ها حساب بانکی داشته باشد رکورد مورد نظر را به all\_banks اضافه کنید ولی اگر یک شخص در چندین بانک حساب داشته باشد در دیکشنری all\_banks مجموع موجودی های شخص در تمام بانک ها را یافته و ثبت کنید.

(۹) رابطه بسط تیلور یک تابع حول نقطه صفر طبق رابطه زیر محاسبه میگردد:

$$f(x) = f(0) + \frac{f'(0)}{1!} x + \frac{f''(0)}{2!} x^2 + \frac{f'''(0)}{3!} x^3 + \dots = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$$

تابع به نام approximate بنویسید که به عنوان ورودی یک چند جمله ای f و نقطه دلخواه x را گرفته و تقریب این چند جمله ای را با n جمله رابطه بالا برگرداند. (۱ نمره)

**راهنمایی:** برای محاسبه مقدار مشتق توابع در پایتون میتونید از کتابخانه sympy استفاده کنید. برای نصب این کتابخانه از دستور pip استفاده کنید.

**!pip install sympy**

در کد زیر نمونه ای از نحوه محاسبه مشتق توابع به کمک این کتابخانه نشان داده شده است.

```
from sympy import Symbol, lambdify
x = Symbol('x')
f_sym = 2 * x **3 + 3 * x + 4
f_prime_sym = f_sym.diff(x)
f = lambdify(x, f_sym)
f_prime = lambdify(x, f_prime_sym)
f(2), f_prime(2)
```

(26, 27)