

بسمه تعالی

موضوع: طراحی text analysis برای سوال ۲ قسمت text

1. استاد خانزادی از ما می‌خواهد برنامه‌ای طراحی کنیم که فایل متنی ارائه شده را از کاربر دریافت کرده و موارد زیر را برای ما مشخص کرده و در خروجی چاپ کند:

- تعداد تکرار هر حرف و عدد در متن
- تعداد واژگان و تعداد تکرار هر کدام
- تعداد جمله‌ها
- تعداد پاراگراف‌ها
- تعداد هر حرف صدadar انگلیسی (a,e,i,o,u)
- میانگین و واریانس تعداد حروف صدadar و بی صدا در هر جمله، هر پاراگراف و در کل متن

همچنین از ما خواسته شده خروجی چاپ شده را در یک فایل متنی و همچنین در یک فایل اکسل ذخیره کنیم. در ضمن می‌توانیم از کتابخانه pandas در پایتون برا فایل اکسل کمک بگیریم و همچنین برای پردازش متن می‌توانیم از کتابخانه های nltk و spacy در پایتون کمک بگیریم.

دید کلی به مساله:

به طور کلی باید یک بستر برای اجرای کد مانند vs code باشد که در آن کد را اجرا کنیم . این کد باید متنی که دریافت میکند را پردازش کند و طبق خواسته های مساله تعیین کند از هر خواسته ، چند مورد در مساله وجود دارد. برای قسمت کد ،می‌توان از زبان های مختلفی استفاده کرد چون سوال مشخص نکرده که چه نوع زبان خاصی مورد استفاده و بحث قرار بگیرد و لب بخ طور کلی ما پایتون رو انتخاب میکنیم برای سهولت و آسانی آن. فایل متنی هم در دسترس ما قرار گرفته است پس باید محیطی را طراحی کنیم که کاربر بتواند فایل متنی خود را در آن آپلود کرده و در نتیجه خروجی کار خود را ببیند و مشاهده کند.

پرسش های کلی:

1. چرا این پروژه برای استاد خانزادی اهمیت دارد؟

پاسخ: به دلایل مختلفی می‌تواند باشد اما دلیل عمده آن ،سنجش توانایی دانشجو و تشخیص توانایی و ارتقا بخشی آن و ...

2. مقدار بودجه و زمان در دسترس چقدر است ؟

پاسخ:مقدار زمان تا روز پنج شنبه مورخ ۱۳ ام اردیبهشت ماه و مقدار بودجه نامشخص میباشد

3. از چه زبان های برنامه نویسی ای میتوان استفاده و بهره برد؟

پاسخ:به طور کلی،زبان خاصی مورد نظر نیست و می‌توان از هر زبان دلخواهی استفاده کرد اما زبان پایتون پیشنهاد می‌شود.

4. کارفرما و مدیر این پروژه کیست؟

پاسخ: استاد پوریا خانزادی

5. مصرف کنندگان و مخاطبان این برنامه چخ کسانی هستند؟

پاسخ: می‌توان هر فردی در هر سنی و جنسیتی باشد
اما اغلب کسانی که می‌توانند یک متن را از نظر
تعداد کلمه، جمله، واژگان خاص و... مورد بررسی
قرار دهد.

6. آیا فقط می‌توان از کلمات انگلیسی استفاده کرد؟

پاسخ: بله، از آنجایی که شروط مساله شامل حروف
صدادار انگلیسی میباشد، میت آن از زبان های
دیگری استفاده کرد اما چندین شرط در آن نقض
می‌شود.

7. نحوه عملکرد این برنامه چگونه میباشد؟

پاسخ: این برنامه ، فایل حاوی کلمات و اعداد مختلف
انگلیسی را از کاربر دریافت کرده و یوسف با توجه
به شرط های ذکر شده، متن را مورد بررسی قرار
داده و تعداد هر شرط را در مقابل آن نمایش میدهد.

☐ موارد کلی به رنگ قرمز

☐ موارد جزئی به رنگ زرد

موارد تحویل دادنی:

1. کد نوشته شده در vs code

2. طراحی و mind map

4. خروجی کد در فایر اکسل

5. خروجی کد در فایل متنی

3. انجام شده برای سوال Text analysis