

Subject

تحریرات مباحث ویرٹن

Year

۱۴۰۴

Month

۱

Day

۲۰

نام مدرس: مباحث ویرٹن

۲ بخش

Advanced Topics

سرما تحریر:

نام دانشجو: معلومہ منیرج

واٹر ٹرم: ورودس بعین ۱۴۰۳

مهندس

نام مدرس: استاد مقید احمد زاده

A. چرا DLT (Distributed Ledger Technology) در مدیریت داده ها کاربرد دارد؟  
 شفافیت و مدیریت مرکز خود، در مدیریت داده ها کاربرد دارد. این فناوری به کاربران اجازه میدهد تا به صورت هر زمان و بدون نیاز به واسطه، اطلاعات را ذخیره و بروز کنند.  
 این ویژگی ها موجب کاهش خطاها و تقلب، افزایش اعتماد و ایمنی در تبادل اطلاعات میشود.

B. چرا برای ذخیره سازی داده ها امن استفاده میشود؟  
 توزیع شده است، برای ذخیره سازی داده ها امن مناسب است. هر بلوک حاوی اطلاعات و هش بلوک قبلی است. به این امر برای جلوگیری از تغییر یا حذف داده ها ضروری است.  
 همچنین استفاده از رمزنگاری در این فناوری امنیت بیشتری را فراهم میکند.  
 C. چرا GANs (Generative Adversarial Networks) در علم داده چیست؟ کاربرد دارند؟

به دلیل توانایی شان برای تولید داده ها جدید و شبیه سازی توزیع های داده ها واقعی بسیار مورد توجه هستند.  
 این شبکه ها میتوانند در زمینه ها مختلفی مانند تولید تصاویر ویدیو، متن و حتی صدا به کار روند و به پژوهشگران کمک کنند تا مدل های پیشرفته تر را برای یادگیری عمیق طراحی کنند.

D. چرا PCA و t-SNE (t-Distributed Stochastic Neighbor Embedding) یک روش غلطی برای کاهش ابعاد است که سعی دارند وابستگی بین متغیرها را در داده ها حفظ کنند، در حالی که t-SNE (t-Distributed Stochastic Neighbor Embedding) سعی دارد وابستگی بین متغیرها را در داده ها حفظ کند.



یک روش غیر خطی است که برای نگه داشتن فاصله های محلی داده ها و به تصویر کشیدن ساختار های پیچیده تر در داده ها طراحی شده است.  $t-SNE$  معمولاً برای داده ها با ابعاد بالا و غیر خطی بهتر عمل میکند.

برای کاهش ابعاد با حفظ واریانس و پیش پردازش داده ها مناسب است.  $PCA$

برای مصور سازی داده ها و کشف ساختار های در داده ها مناسب است.  $t-SNE$

هر دو روش میتوانند مکمل یکدیگر باشند و بسته به نیاز و نوع داده میتوان از هر دو استفاده کرد.

E. چرا UMAP برای Dimensionality Reduction استفاده میشود؟

یک روش کاهش ابعاد است که بر اساس ریاضیات (Approximation and projection)

توپولوژی یک عمل میکند و میتواند ساختار های پیچیده و غیر خطی را به خوبی

حفظ کند. این دارد و برای آنار با داده ها بزرگ  $t-SNE$  روش معمولاً سرعت بالاتر

و کیفیت بهتری مناسب است.

F. چرا Neural Networks برای تحلیل گراف کاربرد دارد؟

به کتابخانه پایتون است که برای ایجاد، تجزیه و تحلیل و مصور سازی گراف ها  $NetworkX$

و شبکه ها طراحی شده است. این کتابخانه ابزار های قدرتمندی برای انجام عملیات بر

روی گراف ها مانند محاسبه ویژگی های ساختاری و الگوریتم های گراف، فراهم

میکند. کاربرد های آن

۱. سهولت استفاده: دارای رابط برنامه نویسی ساده (API) است که ایجاد

ویرایش و تحلیل گراف ها را آسان میکند.



۲. چشمتیابی از انواع گراف ها.

گراف ها غیر جهت دار یا جهت دار می باشند و وزن دار  
مناسب برای مدل سازی شبکه ها اجتماعی، مسیر یابی، سیستم ها  
توصیه گر و غیره ...

۳. الگوریتم ها پیش ساخت

شامل الگوریتم معروف گراف

مانند کوتاه ترین مسیر، جستجوی مولف ها متصل، الگوریتم ها برای

یابی.

کاربردها واقعی

تحلیل شبکه ها اجتماعی (مثلاً شناسایی افراد تأثیرگذار).

مدل سازی زیر ساخت شبکه (اینترنت، حمل و نقل).

سیستم ها توصیه گر (مثل پیدا کردن شباهت بین کاربران).

در مجموعه  $NejWork$  به دلیل سادگی انعطاف پذیر و ابزار ها قدرتمند  
یکی از گزینه ها اصلی برای تحلیل گراف در پایتون است.

۶.  $NLTK$  چرخه برای پردازش زبان طبیعی (NLP) استفاده می شود.

مجموعه ابزار قدرتمند برای پردازش زبان طبیعی است که شامل توابع و

کتابخانه ها مختلف برای تحلیل متن، تجزیه و تحلیل معنایی

و یادگیری ماشین است. این کتابخانه به پژوهشگران و توسعه



دهندگان نگاه میکند تا پروژه های NLP خود را سریع تر و آسان تر انجام دهند.

4. **Scrapy** برای **web scraping** کاربرد دارد. یک چارچوب قدرتمند برای وب اسکرپینگ است که به کاربران اجازه میدهد، راحتی **Scrapy** اطلاعات، از وب سایت ها استخراج کنند.

این ابزار قابلیت های پیشرفته ای مانند مدیریت درخواست ها، مدیریت داده ها و پردازش موازی را ارائه میدهد که باعث میشود فرآیند جمع آوری داده بسیار کارآمدتر باشد.

1. **BeautifulSoup** چارچوبی که برای استفاده میشود.

طراحی XML و HTML یک کتابخانه پایتونی است که برای پارس کردن **BeautifulSoup** شده است.

این کتابخانه به کاربران اجازه میدهد تا به راحتی به عناصر مختلف صفحات وب معمولاً در ترکیب با **BeautifulSoup** دسترسی پیدا کنند و آن را استخراج یا تفسیر دهند. برای وب اسکرپینگ استفاده میشود.