

Subject :

بنا / خرد

Date :

دانشگاه ملی مهارت
آموزشکده میناب

مدرینات بخش فنی (Advanced Topics)

نام و نام خانوادگی: زهرا درخوار، مریم رفیعی زاده، مریم شمسدادی

واحد درسی: مبانی ویر

رشته: مهندسی کامپیوتر

مدرس: آقای محمد احمدزاده

بجای ۱۴۰۳

Subject :

Date :

۱- چرا DLT (Distributed Ledger Technology) در مدیریت داده ها کاربرد دارد ؟

در DLT (Distributed Ledger Technology) به دلیل زیرکاربرد دارد :

۱- شفافیت : دسترسی همگانی به داده ها و تراکنش ها.

۲- امنیت : رمزنگاری و توزیع داده ها برای حفاظت از اطلاعات.

۳- غیر متمرکز : کاهش ریسک های مربوط به نقض امنیت مرکزی.

۴- قابلیت اطمینان : ذخیره سازی داده ها در چندین محل ، افزایش دسترسی پذیری.

۵- کاهش هزینه ها : کاهش نیاز به واسطه ها و کاهش پیچیدگی.

۶- قراردادهای هوشمند : خودکار سازی پروسه ها بدون نیاز به واسطه.

۷- بهبود کارایی : افزایش سرعت و بهینه سازی فرآیندها.

به طور کلی ، DLT به سازمان ها کمک می کند تا داده هایشان را بهتر مدیریت و فرآیندهایشان را بهبود دهند.

Subject:

Date:

۱. Blockchain بسیار برای ذخیره سازی داده ها امن استفاده می شود.

۲. بلاک چین به دلیل امنیتی بالا برای ذخیره سازی داده ها استفاده می شود چون:

۱- رمزنگاری: داده ها با کلیدهای خصوصی و عمومی محافظت می شوند.

۲- توزیع شده: داده ها در شبکه ای از کامپیوترها پخش می شوند.

۳- تغییرناپذیر: تغییر داده ها بسیار دشوار است.

۴- اجماع: اکثریت اعضا باید تراکنش را تأیید کنند.

۵- شفافیت: تمام تراکنش ها قابل مشاهده هستند.

۶- مقایسه در برابر سامانه های سنتز یا سامانه های بسیار دشوار است.

این ویژگی ها بلاک چین را به یک فناوری امن برای ذخیره سازی داده ها تبدیل کرده اند.

Subject :

Date :

۳- میرا (Generative Adversarial Networks) GANs در علم داده بیشتر

کاربرد دارند.

در علم داده (شبکه های مولد تفاضلی) GANs : پی رفته کاربرد دارند چون :

• تولید داده : داده های مصنوعی شبیه واقعی می سازند (مثلاً تصاویر و صدا).

• افزایش داده : تنوع و حجم داده ها را برای آموزش مدل ها زیاد می کنند.

• تنوع و تبدیل : تصاویر را بازنساز می و می توان آن ها را تغییر می دهند.

• تشخیص نا امنی : موارد غیر عادی را در داده ها پیدا می کنند.

• بهبود مدل ها : به عنوان ابزار برای آموزش مدل ها یا دیگر مدل های استفاده می شوند.

به طور خلاصه GANs ابزار قدرتمندی برای تولید، بهبود و تحلیل داده ها در علم داده پی رفته هستند.

۴. PCA و t-SNE چقدر شباهت دارند؟

۱. نوع تکنیک:

PCA: روشی خطی برای کاهش ابعاد داده‌ها.
 t-SNE: روشی غیرخطی که همسایگی‌ها را حفظ می‌کند.

۲. حفظ ساختار:

PCA: ساختار جهانی داده‌ها را بهتر حفظ می‌کند.
 t-SNE: تمرکز بر حفظ ساختار محلی و نهایی خوشه‌ها.

۳. مقیاس پذیری:

PCA: سریع‌تر و مناسب برای داده‌های بزرگ.
 t-SNE: زمان‌برتر و غیرمناسب برای مجموعه‌های بسیار بزرگ.

نقش در بین بردارشی:

PCA: معمولاً برای کاهش ابعاد اولیه استفاده می‌شود.
 t-SNE: بیشتر برای تجسم و کشف الگوهای پیچیده از تحلیل کاربردی دارد.

۴. نتایج بصری:

PCA: نمای خطی و کمتر جذاب.
 t-SNE: تجسم جذاب‌تر و واضح‌تر از خوشه‌ها و الگوها.

★ به طور خلاصه، PCA برای تحلیل‌های آماری و t-SNE برای تجسم خوشه‌ها استفاده می‌شود.

Subject :

Date :

۱- چرا UMAP برای Dimensionality Reduction استفاده می شود؟

UMAP برای کاهش ابعاد استفاده می شود چون :

• ساختار داده ها را خوب حفظ می کند.

• سریع و برای داده های بزرگ مناسب.

• قابل تنظیم و انعطاف پذیر.

• منابع ریاضی قوی دارد.

• می تواند داده های جدید را با همان تبدیل کرد.

Subject:

Date:

۹. NETWORKX چرا برای تحلیل گراف ها کاربرد دارد؟

۱. یک کتابخانه پایتون برای تحلیل NETWORKX گراف ها و شبکه ها است و به دلایل زیر محبوب است:

۱. سادگی و آراستگی: API ساده ای دارد که کار با آن را آسان می کند.

۲. قابلیت های وسیع: شامل الگوریتم های متنوع برای تجزیه و تحلیل گراف ها (مثل جستجوی کوتاه ترین مسیر).

۳. پشتیبانی از انواع گراف: از گراف های جهت دار، بدون جهت و وزن دار پشتیبانی می کند.

۴. تعامل با سایر ابزارها: به راحتی با کتابخانه های دیگری مثل Matplotlib و Pandas ادغام می شود.

۵. تحلیل شبکه های بزرگ: قابلیت کار با شبکه های پیچیده و بزرگ را دارد.

۶. مستندات و جامعه فعال: دارای مستندات جامع و جامعه ای بزرگ برای کمک به کاربران است.

به طور کلی، NETWORKX ابزاری کارآمد برای تحلیل گراف ها است.

Subject :

Date :

✓ NLTK چرا برای پردازش زبان طبیعی (NLP) استفاده می‌شود؟

NLTK به خاطر این ویژگی‌ها متوی NLP پرکاربرده:

• داده‌ها و ابزارهای زیاد: برای یادگیری و تعیین NLP حالت.

• یادگیری آسان: کار کردن با عناصر و دستورات آموخته‌شده زیاد.

• جامع کاربری فعال: اگر سوالی داشته باشید، زود جواب می‌گیرید.

• مناسب: لازم نیست پول بدهید برای استفاده.

Subject:

Date:

۱. چرا scrapy برای web scraping کاربرد دارد؟

web یک فریم ورک قدرتمند برای scrapy در پایتون است که به دلیل ویژگی های scrapy زیرکاربرد است.

۲. سرعت و کارایی: استخراج داده ها را به صورت غیر همزمان و سریع انجام می دهد.

۳. مدیریت آسان درخواست ها: به طور خودکار درخواست ها و پاسخ ها را مدیریت می کند.

۴. ساختار منظم: امکان ساختار منظم داده ها را فراهم می کند.

۵. امکان استخراج داده های پیچیده: برای داده های دینامیک (مانند AJAX) مناسب است.

۶. پشتیبانی از افزونه ها: قابلیت های اضافه برای پردازش و مدیریت داده ها دارد.

۷. مستندات و جامعه فعال: منابع آموزشی و پشتیبانی خوبی دارد.

در کل، scrapy به دلیل قابلیت ها و کارایی بالایی که در زمینه های مختلف برای استخراج داده ها از وب سایت ها است.

Subject:

Date:

4- BeautifulSoup برای کار استفاده می شود.

یک کتابخانه در پایتون است BeautifulSoup که برای استخراج و تجزیه داده ها از فایل های استفاده می شود و ویژگی های XML و HTML کلید آن شامل:

۱- تجزیه آسان و قابلیت تحلیل و استخراج داده ها از ساختارهای HTML و XML به آسان.

۲- سازگاری با HTML ناقص: توانایی کار با صفحات وب که دارای کد ناقص هستند.

۳- انتخاب ساده: استفاده آسان از انتخابگرهای CSS و XPath برای یافتن عناصر.

۴- ویرایش محتوا: امکان تغییر یا حذف عناصر HTML.

۵- سازگاری با کتابخانه های دیگر: می تواند با Requests و Scrapy ترکیب شود.

۶- مستندات خوب: منابع آموزشی و مستندات کامل.

در نتیجه BeautifulSoup ابزاری مناسب برای وب اسکراپینگ و کار با داده های HTML و XML

است.