1. A.چرا (DLT (Technology Ledger Distributed در مدیریت دادهها کاربرد دارد؟

فناوری دفتر کل توزیع‌شده (DLT) به دلایل متعددی در مدیریت داده‌ها کاربرد دارد که مهم‌ترین آن‌ها عبارتند از:

1. تمرکززدایی: DLT داده‌ها را در شبکه‌ای از گره‌ها توزیع می‌کند و نیاز به یک مرجع مرکزی را از بین می‌برد. این امر خطر سانسور، دستکاری و خرابی داده‌ها را کاهش می‌دهد
2. شفافیت: تمام تراکنش‌ها و تغییرات داده‌ها در یک دفتر کل عمومی ثبت می‌شوند که برای همه اعضای شبکه قابل مشاهده است. این امر باعث افزایش اعتماد و مسئولیت‌پذیری می‌شود.
3. امنیت: DLT از الگوریتم‌های رمزنگاری پیشرفته برای محافظت از داده‌ها در برابر دسترسی غیرمجاز و تغییرات استفاده می‌کند. هر تراکنش با یک امضای دیجیتال تأیید می‌شود که جعل آن را غیرممکن می‌سازد.
4. تغییرناپذیری: پس از ثبت یک تراکنش در DLT، نمی‌توان آن را تغییر داد یا حذف کرد. این امر یک سابقه دائمی و قابل اعتماد از داده‌ها ایجاد می‌کند.
5. کارایی: DLT می‌تواند فرایندهای مدیریت داده را ساده‌تر و سریع‌تر کند. به عنوان مثال، در زنجیره تأمین، DLT می‌تواند به ردیابی محصولات از تولید تا مصرف کمک کند و نیاز به واسطه‌ها و کاغذبازی را کاهش دهد.
6. قابلیت همکاری: DLT می‌تواند امکان تبادل داده ایمن و کارآمد بین سازمان‌های مختلف را فراهم کند. این امر به ویژه در صنایعی که نیاز به اشتراک‌گذاری داده‌های حساس دارند، مانند خدمات مالی و مراقبت‌های بهداشتی، مفید است.

به طور خلاصه، DLT یک فناوری قدرتمند است که می‌تواند به بهبود امنیت، شفافیت، کارایی و قابلیت همکاری در مدیریت داده‌ها کمک کند. این امر آن را به یک ابزار ارزشمند برای سازمان‌هایی تبدیل می‌کند که به دنبال راهی برای مدیریت داده‌های خود به طور ایمن، شفاف و کارآمد هستند.

1. B.Blockchainچرا برای ذخیرهسازی دادههای امن استفاده میشود؟

بلاکچین به خاطر امنیت بالا برای ذخیره‌سازی داده‌ها استفاده می‌شود چون:

توزیع‌شده است:اطلاعات در چندین کامپیوتر پخش شده و هک کردن همه آنها تقریبا غیرممکن است.

رمزنگاری قوی: از روش‌های پیچیده ریاضی برای محافظت از اطلاعات استفاده می‌کند.

تغییرناپذیر: بعد از ثبت اطلاعات، تغییر یا حذف آن بسیار دشوار است.

شفافیت (در عین حفظ حریم خصوصی): همه تغییرات قابل ردیابی هستند، اما جزئیات خصوصی می‌توانند محافظت شوند.

C .چرا (Networks Adversarial Generative (GANs در علم داده پیشرفته کاربرد دارند

شبکه‌های مولد تخاصمی (GANs) در علم داده پیشرفته کاربرد دارند به این دلیل که:

تولید داده‌های جدید: GANها می‌توانند داده‌های مصنوعی تولید کنند که شبیه داده‌های واقعی هستند. این برای افزایش حجم داده‌های آموزشی، پر کردن شکاف‌های موجود در داده‌ها یا تولید داده‌هایی که به طور طبیعی در دسترس نیستند، مفید است.

یادگیری بازنمایی: ها می‌توانند بازنمایی‌های پنهانی از داده‌ها را یاد بگیرند. این بازنمایی‌ها می‌توانند برای وظایف دیگری مانند طبقه‌بندی، خوشه‌بندی یا تشخیص ناهنجاری‌ها استفاده شوند.

تبدیل تصویر به تصویر: GANها می‌توانند تصاویر را از یک دامنه به دامنه دیگر تبدیل کنند. این برای وظایفی مانند تبدیل تصاویر سیاه و سفید به رنگی، تبدیل نقاشی‌ها به عکس‌ها یا تولید تصاویر با سبک‌های مختلف مفید است.

بهبود کیفیت تصاویر: GANها می‌توانند تصاویر را با افزایش وضوح، حذف نویز یا ترمیم قسمت‌های آسیب‌دیده بهبود دهند.

به طور خلاصه، GANها ابزاری قدرتمند برای تولید داده‌های جدید، یادگیری بازنمایی و بهبود کیفیت تصاویر هستند که آن‌ها را برای بسیاری از کاربردهای علم داده پیشرفته ارزشمند می‌کند.

:D.SNE-tو PCA چه تفاوتی دارند؟

تفاوت‌های اصلی t-SNE و PCA به صورت خلاصه:

\*pan (تحلیل مولفه‌های اصلی):

روش کاهش ابعاد خطی

حفظ ساختار کلی داده‌ها در ابعاد کمتر (مانند واریانس).

برای تجسم داده‌ها با ابعاد زیاد، اما معمولا خوشه های مجزا را به خوبی نشان نمی دهد.

محاسبه آن سریعتر است.

\*t-SNE (t-distributed Stoch Neighbor Embedding):

روش کاهش ابعاد غیرخطی

حفظ ساختار محلی داده‌ها (نزدیکی نقاط به هم) در ابعاد کمتر.

برای تجسم خوشه‌ها و گروه‌های مجزا در داده‌ها بسیار خوب عمل می‌کند.

محاسبه آن زمان‌برترو پیچیده‌تر است.

به طور خلاصه، PCA برای حفظ ساختار کلی و کاهش ابعاد سریعتر مناسب است، در حالی که t-SNE برای تجسم خوشه‌ها و ساختار محلی پیچیده‌تر داده‌ها بهتر عمل می‌کند اما کندتر است.

E.چرا UMAP برای Reduction Dimensionality استفاده میشود؟

UMAP (Uniform Manifold Approximation and Projection) برای کاهش ابعاد (Dimensionality Reduction) استفاده می‌شود چون:

حفظ ساختار کلی و محلی: UMAP سعی می‌کند هم روابط بزرگ‌مقیاس (ساختار کلی) و هم روابط کوچک‌مقیاس (ساختار محلی) داده‌ها را در فضای با ابعاد کمتر حفظ کند. این باعث می‌شود در مقایسه با PCA که بیشتر بر ساختار کلی تمرکز دارد و t-SNE که بیشتر بر ساختار محلی تاکید دارد، تعادل بهتری برقرار کند.

سرعت و مقیاس‌پذیری: UMAP معمولاً سریعتر از t-SNE است و می‌تواند به خوبی با مجموعه‌های داده بزرگ مقیاس‌پذیر باشد.

تئوری قوی: UMAP بر پایه مفاهیم ریاضیاتی نسبتاً قوی‌تری نسبت به t-SNE بنا شده است.

به طور خلاصه، UMAP انتخاب خوبی برای کاهش ابعاد است به دلیل توانایی آن در حفظ ساختار کلی و محلی داده‌ها، سرعت بالا، و مقیاس‌پذیری خوب.

F.NetworkXچرا برای تحلیل گرافها کاربرد دارد؟

NetworkX برای تحلیل گراف‌ها کاربرد دارد چون:

ساختارهای داده‌ای منعطف: امکان ساخت و دستکاری آسان گراف‌ها و شبکه‌ها با انواع گره‌ها و یال‌ها را فراهم می‌کند.

الگوریتم‌های متنوع: شامل مجموعه‌ای گسترده از الگوریتم‌های تحلیل گراف (مانند کوتاه‌ترین مسیر، مرکزیت، تشخیص اجتماع و غیره) است.

ابزارهای تجسم: امکان نمایش گراف‌ها را به صورت بصری فراهم می‌کند.

جامعه کاربری بزرگ: مستندات خوب و پشتیبانی گسترده‌ای دارد.

به‌طور خلاصه، NetworkX ابزاری قدرتمند و انعطاف‌پذیر برای ایجاد، تحلیل و تجسم گراف‌ها و شبکه‌ها است.

J.NLTKچرا برای پردازش زبان طبیعی (NLP (استفاده میشود؟

NLTK برای پردازش زبان طبیعی (NLP) استفاده می‌شود چون:

مجموعه ابزار جامع: شامل ابزارهای متنوعی برای توکن‌بندی، ریشه‌یابی، برچسب‌گذاری اجزای سخن، تجزیه نحوی، و غیره است.

منابع داده غنی: دارای مجموعه‌های داده (corpora) و لغت‌نامه‌های از پیش آماده شده است.

یادگیری آسان: رابط کاربری ساده و آموزشی دارد.

جامعه فعال: مستندات خوب و پشتیبانی گسترده‌ای دارد.

به طور خلاصه، NLTK یک پکیج قدرتمند و آموزشی برای شروع و انجام پروژه‌های NLP است.

H.چرا Scrapy برای Scraping Web کاربرد دارد؟

Scrapy برای Web Scraping کاربرد دارد چون:

فریم‌ورک قدرتمند: یک فریم‌ورک کامل برای خزیدن و استخراج داده از وب‌سایت‌ها است.

عملکرد بالا: طراحی شده برای مدیریت درخواست‌های متعدد و استخراج سریع داده‌ها.

قابلیت گسترش: امکان اضافه کردن میان‌افزارها و پایپ‌لاین‌های سفارشی برای پردازش داده‌ها.

مدیریت آسان: ابزارهایی برای مدیریت خزش‌ها و ذخیره‌سازی داده‌ها فراهم می‌کند.

به طور خلاصه، Scrapy یک فریم‌ورک کارآمد و قابل تنظیم برای خودکارسازی فرآیند استخراج داده از وب است.

1. P.BeautifulSoupچرا برای کاری استفاده می شود؟

Beautiful Soup برای تحلیل و پیمایش آسان اسناد HTML و XML استفاده می‌شود. به عبارت دیگر، به شما کمک می‌کند تا به راحتی اطلاعات مورد نظرتان را از صفحات وب استخراج کنید.