



دانشکده ملی مهارت

شهرستان میناب

تمرین های درس مباحث ویژه(بخش ششم)

مرتبط به درس:

مباحث ویژه

گردآورنده:

محمدحسین خبیری

استاد:

محمد احمدزاده

ترم تحصیلی بهمن ۱۴۰۳-۱۴۰۴

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

**Distributed Ledger Technology (DLT)** در مدیریت داده‌ها کاربرد دارد چون:

* **توزیع‌شده و غیرمتمرکز**: داده‌ها در شبکه‌ای از گره‌ها ذخیره می‌شوند و هیچ نقطه مرکزی برای کنترل وجود ندارد، که امنیت و شفافیت را افزایش می‌دهد.
* **امنیت بالا**: با استفاده از رمزنگاری و مکانیزم‌های اجماع، امکان تغییر یا دستکاری داده‌ها بسیار کم است.
* **شفافیت و قابلیت ردیابی**: تمام تراکنش‌ها و تغییرات در دفتر ثبت شده و قابل پیگیری هستند.
* **کاهش هزینه‌ها**: DLT می‌تواند نیاز به واسطه‌ها را حذف کند و هزینه‌های مربوط به تأمین امنیت و پردازش داده‌ها را کاهش دهد.

1. Blockchain چرا برای ذخیره‌سازی داده‌های امن استفاده می‌شود؟

* Blockchain برای ذخیره‌سازی داده‌های امن استفاده می‌شود چون:
* غیرقابل تغییر بودن: پس از ثبت یک داده در بلاکچین، تغییر آن بسیار دشوار است زیرا هر بلاک به بلاک قبلی متصل است و برای تغییر یک بلاک باید همه بلاک‌های بعدی تغییر کنند.
* رمزنگاری: داده‌ها با استفاده از الگوریتم‌های رمزنگاری قوی محافظت می‌شوند که دسترسی غیرمجاز به اطلاعات را غیرممکن می‌کند.
* توزیع‌شده بودن: داده‌ها در شبکه‌ای از گره‌ها ذخیره می‌شوند، بنابراین حتی اگر یک یا چند گره هک شوند، داده‌ها همچنان امن باقی می‌مانند.
* شفافیت و پیگیری: تمام تراکنش‌ها در دفترکل عمومی ثبت می‌شوند و هر تغییر قابل پیگیری است.

1. چرا GANs (Generative Adversarial Networks) در علم داده پیشرفته کاربرد دارند؟

* GANs (Generative Adversarial Networks) در علم داده پیشرفته کاربرد دارند چون:
* تولید داده‌های جدید: GANها قادر به تولید داده‌های جدید مشابه داده‌های واقعی هستند که در حوزه‌هایی مانند تصویرسازی، تولید ویدئو، و حتی تولید متن کاربرد دارند.
* یادگیری ویژگی‌های پیچیده: GANها می‌توانند ویژگی‌های پیچیده و غیرخطی داده‌ها را یاد بگیرند و این امکان را برای حل مسائل پیچیده فراهم کنند.
* ارتقاء کیفیت داده‌ها: از GANها برای بهبود کیفیت داده‌های ناقص یا خراب، مانند بازسازی تصاویر ویدئویی، استفاده می‌شود.
* نوآوری در مدل‌های هوش مصنوعی: GANها در تحقیق و توسعه مدل‌های هوش مصنوعی کاربرد دارند، به ویژه در ایجاد محتواهای جدید یا پیش‌بینی‌های دقیق‌تر.

1. t-SNE و PCA چه تفاوتی دارند؟

* t-SNE و PCA هر دو برای کاهش ابعاد داده‌ها استفاده می‌شوند، اما تفاوت‌های اصلی دارند:
* PCA (Principal Component Analysis):
* خطی: PCA برای کاهش ابعاد از ترکیب خطی ویژگی‌ها استفاده می‌کند.
* هدف: بیشترین واریانس داده‌ها را در ابعاد جدید حفظ می‌کند.
* کاربرد: برای داده‌های با ارتباطات خطی مناسب است.
* t-SNE (t-Distributed Stochastic Neighbor Embedding):
* غیرخطی: t-SNE برای کاهش ابعاد از روش‌های غیرخطی استفاده می‌کند.
* هدف: ساختار همسایگی داده‌ها را در ابعاد پایین حفظ می‌کند.
* کاربرد: برای داده‌های پیچیده و غیرخطی مانند داده‌های تصویری یا متنی مفید است.

1. چرا UMAP برای Dimensionality Reduction استفاده می‌شود؟

* UMAP (Uniform Manifold Approximation and Projection) برای کاهش ابعاد استفاده می‌شود چون:
* دقت بالا در حفظ ساختار داده‌ها: UMAP قادر است ساختارهای پیچیده و غیرخطی داده‌ها را در ابعاد پایین به خوبی حفظ کند.
* سرعت بیشتر: نسبت به t-SNE، UMAP سرعت بالاتری دارد و می‌تواند داده‌های بزرگتری را سریع‌تر پردازش کند.
* قابلیت تفسیرپذیری بهتر: UMAP قابلیت حفظ روابط همسایگی و ایجاد نمایش‌های با معنی در ابعاد پایین‌تر را دارد.
* انعطاف‌پذیری بالا: می‌تواند در انواع مختلف داده‌ها (مانند تصاویر، متون و داده‌های عددی) به خوبی عمل کند.

1. NetworkX چرا برای تحلیل گراف‌ها کاربرد دارد؟

* NetworkX برای تحلیل گراف‌ها کاربرد دارد چون:
* مدیریت گراف‌ها: NetworkX امکان ایجاد، تغییر و تجزیه و تحلیل انواع مختلف گراف‌ها (جهت‌دار، بدون جهت، وزنی و غیره) را فراهم می‌کند.
* الگوریتم‌های گرافی: این کتابخانه مجموعه‌ای از الگوریتم‌های پرکاربرد برای تحلیل گراف‌ها مانند جستجوی عمق و عرض، یافتن کوتاه‌ترین مسیر، تشخیص اجتماع‌ها و غیره را ارائه می‌دهد.
* تجزیه و تحلیل شبکه‌ها: NetworkX قابلیت تحلیل شبکه‌ها و روابط پیچیده بین عناصر را دارد که برای مسائل شبکه‌های اجتماعی، مسیرهای حمل و نقل، و حتی شبکه‌های عصبی مفید است.
* تجسم گراف‌ها: این کتابخانه ابزارهایی برای رسم گراف‌ها و نمایش بصری روابط میان گره‌ها فراهم می‌کند.

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

* NLTK (Natural Language Toolkit) برای پردازش زبان طبیعی (NLP) استفاده می‌شود چون:
* ابزارهای متنوع: NLTK مجموعه‌ای گسترده از ابزارها برای پردازش و تحلیل متون طبیعی مانند توکن‌سازی، برچسب‌گذاری بخش‌های سخن، تحلیل نحوی و معنایی و غیره را فراهم می‌کند.
* دسترس‌پذیری داده‌ها: این کتابخانه شامل مجموعه‌های داده‌ای بزرگ برای آموزش مدل‌های NLP مانند متون خبری، ادبیات و غیره است.
* پشتیبانی از الگوریتم‌های مختلف: NLTK از الگوریتم‌ها و تکنیک‌های مختلف یادگیری ماشین و پردازش زبان طبیعی پشتیبانی می‌کند، مانند تحلیل احساسات، استخراج ویژگی‌ها و شبیه‌سازی زبان.
* سادگی استفاده: با توجه به مستندات و رابط کاربری ساده، NLTK برای محققان، توسعه‌دهندگان و دانشجویان مناسب است.

1. چرا Scrapy برای Web Scraping کاربرد دارد؟

* Scrapy برای Web Scraping کاربرد دارد چون:
* سرعت بالا: Scrapy طراحی شده است تا به صورت موازی و سریع داده‌ها را از وب‌سایت‌ها استخراج کند.
* ساختاردهی آسان داده‌ها: این ابزار امکان ساختاردهی داده‌ها به فرمت‌های مختلف مانند JSON، CSV، و XML را به راحتی فراهم می‌کند.
* پشتیبانی از درخواست‌های پیچیده: Scrapy از درخواست‌های HTTP، مدیریت کوکی‌ها، دنبال کردن لینک‌ها و حتی پردازش محتوای JavaScript پشتیبانی می‌کند.
* توانمندی در استخراج داده‌های پیچیده: با استفاده از CSS selectors و XPath می‌توان به راحتی داده‌های خاص و پیچیده را از صفحات وب استخراج کرد.
* مدیریت Crawl و Scrape: Scrapy ابزارهایی برای مدیریت crawl و scrape به صورت مؤثر، از جمله محدود کردن سرعت درخواست‌ها و جلوگیری از بلاک شدن توسط سایت‌ها، فراهم می‌کند.

1. BeautifulSoup چرا برای کاری استفاده می شود؟

* BeautifulSoup برای Web Scraping استفاده می‌شود چون:
* سادگی و قابلیت استفاده: این کتابخانه برای تجزیه و تحلیل و استخراج داده‌ها از HTML و XML بسیار ساده و کاربرپسند است.
* پشتیبانی از انواع مختلف HTML: BeautifulSoup قادر است با HTML‌های ناقص یا خراب کار کند و آن‌ها را به طور صحیح پردازش کند.
* جستجوی مؤثر در سند: این ابزار امکان جستجو و انتخاب عناصر خاص در صفحات وب از طریق CSS selectors یا ویژگی‌های HTML مانند کلاس‌ها و شناسه‌ها را فراهم می‌کند.
* سازگاری با سایر کتابخانه‌ها: BeautifulSoup به خوبی با کتابخانه‌هایی مانند Requests یا Selenium ترکیب می‌شود تا درخواست‌های وب را ارسال کرده و داده‌ها را استخراج کند.