

دوره یادگیری ماشین

تمرین ۳

Mall Customer Segmentation Data

مقدمه

بررسی رفتار و ترجیحات مشتریان برای کسب و کارها بسیار حیاتی است تا استراتژی‌های بازاریابی خود را سفارشی‌سازی کرده و رضایت مشتری را افزایش دهند. در این تمرین، به بررسی مجموعه داده تقسیم بندی مشتریان مرکز خرید می‌پردازیم، یک منبع ارزشمند که شامل اطلاعات دموگرافیکی و هزینه‌های مشتریان مرکز خرید است. با استفاده از تکنیک‌های خوشه‌بندی بر روی این مجموعه داده، هدف ما از بین بردن بخش‌های مختلف مشتریان بر اساس الگوهای خرید آنها است.

شناخته شدن تقسیم بندی مشتریان به کسب و کارها امکان سفارشی‌سازی کمپین‌های بازاریابی خود، بهینه‌سازی پیشنهادات محصول و بهبود تجربه کلی مشتری را می‌دهد. از طریق این تجزیه و تحلیل، ما سعی می‌کنیم به دستاوردهایی که می‌تواند تصمیم‌گیری‌های استراتژیک برای مدیریت مرکز خرید را رهنموده، دست یابیم.

اهداف

- استفاده از تکنیک‌های خوشه‌بندی برای تقسیم بندی مشتریان بر اساس ویژگی‌های دموگرافیکی و رفتار خرید آنها
- بررسی مجموعه داده تقسیم بندی مشتریان مرکز خرید برای به دست آوردن بینش در ترجیحات و ویژگی‌های مشتریان
- اعمال تکنیک‌های پیش‌پردازش داده برای اطمینان از کیفیت و سازگاری مجموعه داده برای تحلیل خوشه‌بندی
- پیاده‌سازی و مقایسه الگوریتم‌های مختلف خوشه‌بندی برای شناسایی رویکرد مناسب‌تر برای تقسیم بندی مشتریان
- تفسیر خوشه‌های حاصل و استخراج بینش‌های معنادار در مورد بخش‌های مشتری و تأثیرات بازاریابی مربوطه
- ارائه پیشنهادات به مدیریت مرکز خرید بر اساس نتایج خوشه‌بندی برای بهینه‌سازی استراتژی‌های بازاریابی و افزایش رضایت مشتری

توضیحات درباره مجموعه داده

مجموعه داده شامل ویژگی‌های زیر است:

- **CustomerID**: شناسه یکتای هر مشتری
- **Gender**: جنسیت مشتری (مرد/زن)
- **Age**: سن مشتری
- **Annual Income**: درآمد سالانه مشتری به هزار دلار
- **Spending Score**: امتیاز اختصاصی به مشتری بر اساس رفتار خرید و تعاملات گذشته آنها

وظایف

1. بررسی داده

- a. بارگذاری مجموعه داده و انجام بررسی اولیه داده
- b. بررسی خلاصه‌ی آماری مجموعه داده
- c. بررسی مقادیر گم‌شده و اگر لازم با آن‌ها برخورد مناسب انجام دهید
- d. بصری‌سازی توزیع ویژگی‌ها با استفاده از نمودارهای مناسب (مانند هیستوگرام، نمودار جعبه)

2. انتخاب ویژگی‌ها

- a. تصمیم بگیرید که کدام ویژگی‌ها برای خوشه‌بندی (مانند سن، درآمد سالانه، امتیاز هزینه) استفاده شود.
- b. انتخاب خود را براساس ویژگی‌های مجموعه داده و اهداف تقسیم بندی مشتریان خود را توجیه کنید.

3. پیش‌پردازش داده

- a. ویژگی‌های انتخاب شده را به منظور اطمینان از هم‌مقیاس بودن آن‌ها، نرمال کنید.
- b. اهمیت نرمال‌سازی را در زمینه خوشه‌بندی توضیح دهید.

4. الگوریتم های خوشه‌بندی

- a. حداقل دو الگوریتم خوشه‌بندی (مانند K-means، خوشه‌بندی سلسله مراتبی) را انتخاب کنید تا مشتریان را تقسیم کنید.
- b. الگوریتم‌های انتخاب شده را با استفاده از کتابخانه‌های مناسب پیاده‌سازی کنید (مانند scikit-learn در پایتون).
- c. پارامترها را اگر لازم است تنظیم کنید.
- d. توضیحی درباره کلیت هر الگوریتم خوشه‌بندی و نحوه کار آنها ارائه دهید.

5. تجزیه و تحلیل خوشه‌بندی

- a. الگوریتم‌های خوشه‌بندی را بر داده‌های پیش‌پردازش شده اعمال کنید.
- b. خوشه‌ها را با استفاده از نمودارهای پراکندگی یا تکنیک‌های بصری مناسب دیگر بصری‌سازی کنید.
- c. خوشه‌های حاصل را بر اساس ویژگی‌های آن‌ها تحلیل و تفسیر کنید.
- d. استراتژی‌های بازاریابی ممکن برای هر بخش مشتری را بررسی کنید.

6. ارزیابی

- a. کیفیت خوشه‌ها را با استفاده از معیارهای اعتباری داخلی (مانند امتیاز سیلونت) ارزیابی کنید.
- b. عملکرد الگوریتم‌های مختلف خوشه‌بندی را مقایسه کنید.
- c. بینش‌هایی در مورد نقاط قوت و ضعف هر الگوریتم در این زمینه ارائه دهید.

7. نتیجه گیری

- a. نتایج کلیدی تحلیل خود را خلاصه کنید.
- b. درباره اهمیت تقسیم بندی مشتریان برای تصمیم‌گیری‌های کسب و کاری توضیح دهید.
- c. پیشنهاداتی برای مدیریت مرکز خرید بر اساس نتایج خوشه‌بندی را ارائه دهید.

مجموعه داده

شما می‌توانید مجموعه داده مورد نیاز این تمرین را از این [لینک](#) دانلود کنید.

منابع

- [Sklearn Clustering Documentation](#)