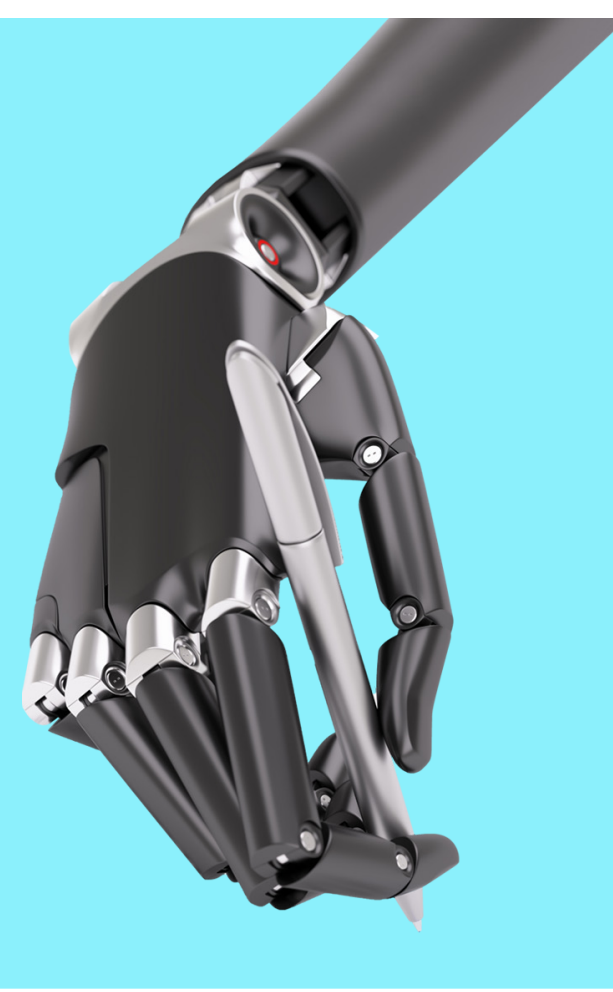


چالش ورودی بوت کمپ ماشین لرنینگ

MACHINE LEARNING BOOTCAMP



چالش ورودی بوت کمپ ماشین لرنینگ

از اینکه علاقه مند به شرکت در بوت کمپ ماشین لرنینگ رهنماکالج هستید، خوشحالیم. در فرآیند ثبت نام علاوه بر سوالاتی که در مورد رزومه و انگیزه شما پرسیده شده، لازم است به سوالات زیر پاسخ دهید و پاسخ تان را در فرم ثبت نام برای ما بازگذاری کنید. برای پاسخ به سوالات به داده هایی نیاز دارید که می توانید آنها را در [اینجا](#) ببینید.

همچنین توجه داشته باشید که در مسیر حل تمرین ها، مهم ترین چیز تلاش برای یادگیری و درگیر شدن با مسئله ست. اگر در بخشی از تمرین احساس کردید موضوع برایتان چالش برانگیز است یا خارج از توان فعلیتان به نظر می رسد، اصلاً نگران نباشید؛ این بخشی طبیعی از فرایند یادگیریست. ما در این دوره بیشتر از هر چیز، روی همین تلاش و پشتکار تمرکز داریم و همراهتان هستیم تا مسیر یادگیری برایتان هموارتر شود. پس با خیال راحت جلو بروید—در ارزیابی تمرین ها هم این موضوع در نظر گرفته می شود.

«رمزگشایی از الگوهای پنهان»

زمینه: تصور کنید که به عنوان یک دانشمند داده در یک مؤسسه تحقیقاتی مشغول به کار هستید و مجموعه ای از داده های چندبعدی و بدون برچسب در اختیار دارید. هدف شما کشف ساختارهای پنهان و الگوهای موجود در این داده ها ست.

۱. کاوش اولیه داده ها

- **بارگذاری و بررسی:** مجموعه داده را بارگذاری کرده و ساختار آن را بررسی کنید. چند بعد دارد؟ آیا داده های گمشده وجود دارد؟
- **بصری سازی:** از تکنیک های کاهش ابعاد مانند PCA یا t-SNE برای نمایش داده ها در فضای ۲ یا ۳ بعدی استفاده کنید. چه الگوهایی مشاهده می کنید؟
- **تحلیل آماری:** میانگین، میانه، و انحراف معیار هر ویژگی را محاسبه کنید. آیا ویژگی هایی با توزیع غیرمعمول وجود دارد؟
- **بررسی همبستگی:** ماتریس همبستگی بین ویژگی ها را محاسبه و تحلیل کنید. کدام ویژگی ها بیشترین همبستگی را دارند؟

۲. تحلیل خوشه بندی

- **انتخاب الگوریتم:** الگوریتم های مختلف خوشه بندی مانند K-Means، DBSCAN، و خوشه بندی سلسله مراتبی را اعمال کنید.
- **ارزیابی:** از معیارهای داخلی مانند امتیاز Silhouette یا شاخص Davies-Bouldin برای ارزیابی کیفیت خوشه ها استفاده کنید.
- **مقایسه:** نتایج الگوریتم های مختلف را مقایسه کنید. کدام الگوریتم بهترین عملکرد را داشت و چرا؟
- **تحلیل حساسیت:** تأثیر تغییر پارامترهای الگوریتم ها (مانند تعداد خوشه ها در K-Means) را بررسی کنید.

۳. تشخیص الگوهای پیشرفته

- **کاهش ابعاد پیشرفته:** از تکنیک های پیشرفته تری مانند UMAP برای کاهش ابعاد استفاده کنید و نتایج را با PCA و t-SNE مقایسه کنید.
 - **تشخیص ناهنجاری:** آیا نقاط پرت یا ناهنجاری در داده ها وجود دارد؟ از چه روش هایی برای شناسایی آن ها استفاده می کنید؟
 - **توصیف خوشه ها:** ویژگی های متمایز هر خوشه را شناسایی و توصیف کنید. آیا الگوهای خاصی مشاهده می کنید؟
 - **تحلیل پایداری:** پایداری خوشه ها را با استفاده از تکنیک هایی مانند خوشه بندی مجدد با داده های نمونه برداری شده بررسی کنید.
- #### ۴. تحقیق و فرضیه سازی
- **چالش های یادگیری بدون نظارت:** در مورد چالش های مرتبط با خوشه بندی در فضاهای با ابعاد بالا تحقیق کنید. این چالش ها چگونه بر این مجموعه داده تأثیر می گذارند؟
 - **تأثیر کاهش ابعاد:** بررسی کنید که چگونه تکنیک های مختلف کاهش ابعاد بر نتایج خوشه بندی تأثیر می گذارند.
 - **کاربردهای دنیای واقعی:** سناریوهای واقعی را پیشنهاد دهید که در آن ها تحلیل چنین داده های پیچیده و بدون برچسب ضروری است.
 - **پیشنهاد بهبود:** بر اساس تحلیل های خود، پیشنهادهایی برای بهبود الگوریتم های خوشه بندی ارائه دهید.

موارد قابل ارائه

گزارش: گزارشی جامع شامل روش شناسی، یافته ها، و تفسیرهای شما. نمودارها و توجیه انتخاب الگوریتم ها و پارامترها را درج کنید.

ارائه: یک ارائه خلاصه از رویکرد و نکات کلیدی خود تهیه کنید. آماده باشید تا در مورد چالش ها و نحوه مواجهه با آن ها بحث کنید.

مخزن کد: مخزنی از کدهای مستند با دستورالعمل های واضح برای بازتولید تحلیل های شما ارائه دهید.

مطالعه کتاب “پنج نقص عملکرد تیم”

معرفی کتاب: کتاب پنج نقص عملکرد تیم نوشته ی پاتریک لنچونی یکی از پرفروش ترین کتاب های مدیریتی و رهبری در جهان است که با زبانی ساده، در قالب یک داستان واقعی، مشکلات اساسی تیم ها را بیان می کند.

۱. نسخه متنی یا صوتی کتاب را مطالعه کنید (نسخه صوتی و متنی در پلتفرم های مختلف موجود است).

۲. تحلیل شخصی:

- کدام یک از این نقص ها را در تیم هایی که تا به حال با آن ها کار کرده اید، تجربه کرده اید؟
- فکر می کنید کدام نقص بیشترین تأثیر را روی شکست تیم ها دارد؟ چرا؟
- اگر شما رهبر تیم باشید، چگونه سعی می کنید این مشکلات را حل کنید؟

۳. گزارش کوتاه:

گزارشی یک صفحه ای درباره آموخته های خود از کتاب تهیه کرده و مثال هایی واقعی از تجربه های خود یا تیم های دیگر ارائه دهید.

هدف نهایی تمرین:

با این تمرین می خواهیم شما نه تنها در زمینه الگوریتم ها و مدل های یادگیری ماشین متخصص شوید، بلکه مهارت های مهمی مانند کار تیمی، اعتماد سازی، و رهبری مؤثر را نیز تقویت کنید، چون موفقیت یک پروژه فنی، فقط به کدنویسی وابسته نیست.

علاوه بر چالش اصلی تعدادی سوال امتیازی برای شما در نظر گرفته شده که با پاسخ به آنها می توانید شانس خود را برای شرکت در دوره افزایش دهید برای دیدن سوالات امتیازی [اینجا](#) کلیک کنید.

دقت کنید!

- لطفا در هر سوال که با کمبود اطلاعات مواجه بودید، فرضیاتی را در نظر بگیرید و این فرضیات را در آخر پاسخ تان ذکر کنید.
- پاسخ های خود را باید در قالب یک فایل PDF ارسال کنید که نام آن ترکیبی از نام و نام خانوادگی شما به طور مثال، Ali.Akbari باشد.
- هرچند جست و جوی فردی بخشی از مراحل حل مسئله است با این وجود نهایتاً ما در ارزیابی ها و بررسی پاسخ ها مبنای بر صداقت شما می گذاریم و فرض می کنیم به این موضوع واقف هستید که پاسخ غیرصادقانه شما، ما را در ارزیابی صحیح دچار گمراهی می کند و ممکن است وارد دوره ای شوید که متناسب با نیازتان نیست.
- با توجه به محدود بودن ظرفیت دوره، چالش ها به ترتیب ارسال، بررسی و ارزیابی می شوند و با پر شدن ظرفیت، مهلت ارسال چالش به اتمام می رسد.

به امید دیدارتان هستیم