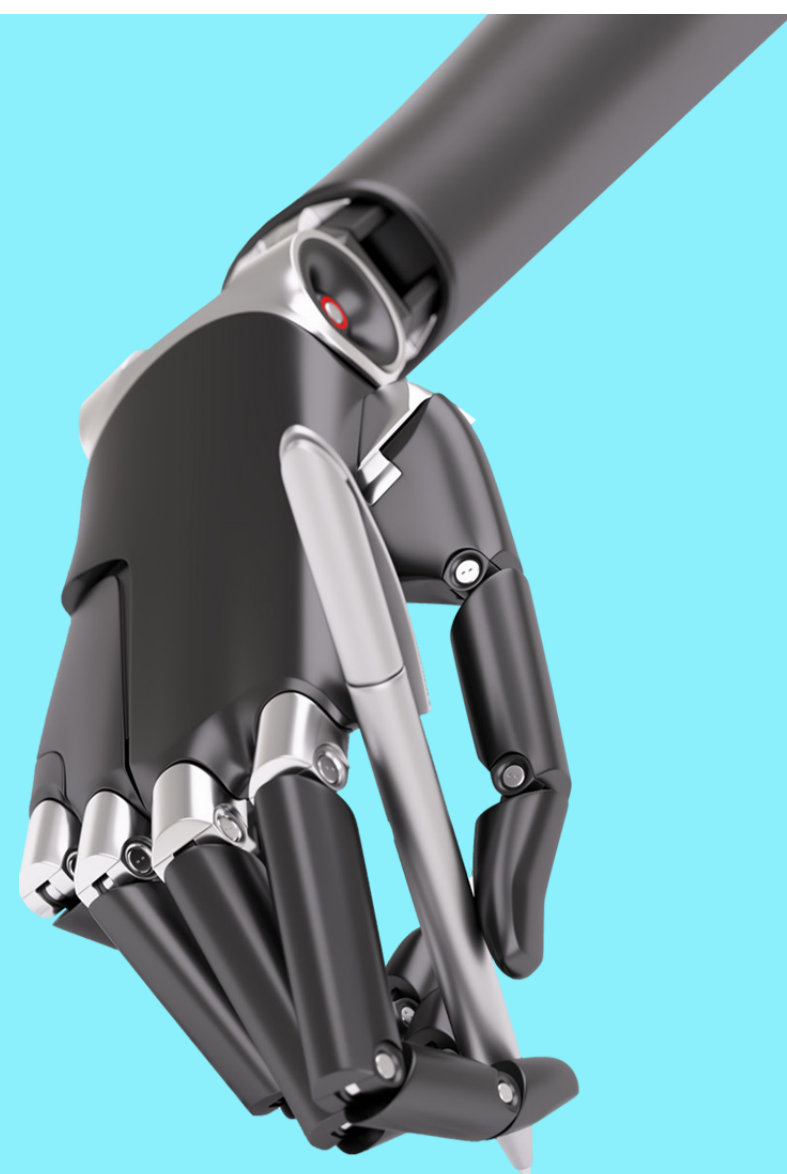


چالش ورودی بوت کمپ

# یادگیری ماشین

## MACHINE LEARNING BOOTCAMP



### چالش ورودی بوت کمپ یادگیری ماشین

از اینکه علاقمند به شرکت در بوت کمپ یادگیری ماشین رهنما کالج 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 پاسختان را در فرم ثبت نام برای ما بارگذاری کنید.

**سوال اول:** سوال اول: شما به عنوان دانشمند داده در یک شرکت طراحی فضای سبز استخدام شده‌اید. شرکت شما از تصاویر هوایی استفاده کرده تا بتواند نقاطی که فضای سبز خوبی ندارند را مشخص کند.

در دیتاست، مختصات  $x$  این تصاویر با `feature_1` و مختصات  $y$  آن‌ها با `feature_2` نام گذاری شده است. شرکت قصد دارد ۵ تیم اعزام کند. تسک شما این است که به هر نقطه یک تیم را اختصاص دهید به گونه‌ای که هزینه رفت و آمد اعضای تیم حداقل شود. (برای مثال منطقی نیست نقطه شمال شهر و جنوب شهر را به تیم شماره یک اختصاص دهید چون فاصله زیادی از هم دارند و بهینه نیست.) داخل دادگان یک فیچر `team_number` اضافه کنید و شماره تیم مورد نظر را وارد کنید و همراه با موارد خواسته شده زیر ارسال کنید.

۱- یک متن شامل روش حل مسئله.

(توضیحات ارسالی خود را در حداقل ۲ پاراگراف و روش حل را به صورت کامل توضیح دهید.)

۲- کد پایتونی که با آن مسئله را حل کرده‌اید به همراه `Documentation` مناسب.

(کد باید به گونه‌ای پیاده‌سازی شده باشد که قابل فهم باشد و در صورت نیاز توضیحاتی در آن ارائه شود.)

۳- مصورسازی گروه‌بندی‌ها بر روی نقشه به کمک `Matplotlib` و ارائه آن در فایل خروجی امتیاز مثبت دارد.

**سوال دوم:** کمند به تازگی در یک آژانس تبلیغاتی به عنوان کارآموز یادگیری ماشین استخدام شده است. از او خواسته شده است تا به کمک نمونه‌های جمع‌آوری شده از تجربیات قبلی شرکت مدلی بسازد تا بتواند نمونه‌های موفق در ارائه تبلیغات به کاربران را از نمونه‌های ناموفق جدا کند. اما متأسفانه این نمونه‌ها و ویژگی‌های مرتبط با آن‌ها با دقت کافی جمع‌آوری نشده‌اند. او اکنون از شما کمک می‌خواهد تا بتواند به سوالات زیر پاسخ دهد.

۱- ویژگی‌های تاثیرگذار برای جداسازی نمونه‌های موفق از ناموفق کدامند؟ (روش خود را کامل توضیح دهید و از مصورسازی استفاده کنید.)

۲- چه معیار ارزیابی برای این دسته‌بندی مناسب است؟

۳- کدام مدل مناسب برای چنین دسته‌بندی است؟ چرا؟ (علت انتخاب خود را در یک پاراگراف توضیح دهید.)

### دقت کنید!

• ضروری است کد ارسالی با زبان پایتون باشد و برای این سوال فقط از کتابخانه‌های `pandas`, `matplotlib`, `NumPy` استفاده شود و از کتابخانه `Scikit-learn` در حل این چالش استفاده نشود.

• پاسخ‌های خود را باید در قالب یک فایل `ZIP` ارسال کنید که نام آن ترکیبی از نام و نام خانوادگی شما به طور مثال، `Ali.Akbari` باشد.

• هرچند جست‌وجوی فردی بخشی از مراحل حل مسئله است با این وجود نهایتاً ما در ارزیابی‌ها و بررسی پاسخ‌ها مبنای بر صداقت شما می‌گذاریم و فرض می‌کنیم به این موضوع واقف هستید که پاسخ غیرصادقانه شما، ما را در ارزیابی صحیح دچار گمراهی می‌کند و ممکن است وارد دوره‌ای شوید که متناسب با نیازتان نیست.

به امید دیدارتان هستیم