



در این پروژه از شما می‌خواهیم تا طبقه‌بندهای مختلف را بر روی دیتاست مورد نظر آموزش دهید و آن را ارزیابی کنید تا با متوذهای ensemble آشنا شوید. نکات قابل توجه:

- در تمام فازها دقت train و test مدل را گزارش کنید.
- در تمام فازها معیارهای precision و recall و f1\_score را هم برای مجموعه آموزشی و هم برای مجموعه تست محاسبه کنید.
- به یاد داشته باشید که لازم است نمودار دقت بر اساس تعداد Estimator ها را برای هر کدام از الگوریتم‌های Random Forest، Bagging، Adaboost ترسیم کنید.
- برای تمام مدل‌ها hyperparameter tuning انجام دهید تا به بهترین دقت برسید.
- تمام خروجی‌ها را در داکيومنت خود تحلیل و بررسی کنید.

## دیتاست

این دیتاست که شامل 30000 داده‌ی دو بعدی است و دو کلاس دارد؛ در قالب دو فایل train.csv و test.csv بر روی ویو آپلود شده است. در ابتدا با ویژوال کردن داده‌های آموزشی به درکی از آن‌ها برسید.

## فاز اول: Bagging

در این فاز از شما می‌خواهیم تا الگوریتم Bagging را که در درس یادگرفته‌اید پیاده‌سازی کنید. توجه کنید که برای این الگوریتم حق استفاده از کتابخانه و توابع آماده را ندارید و صرفاً برای ساخت درخت تصمیم‌های مورد نیاز می‌توانید از کتابخانه sklearn استفاده کنید.

• در ادامه خروجی پیش‌بینی شده الگوریتم Bagging و ۵ کلسیفایر Decision Tree را بر روی دیتاست Visual کرده و برداشت خود از عملکرد هر کدام از کلسیفایرهای ۶ تصویر را در فایل گزارش خود بنویسید (سعی کنید درختانی را انتخاب کنید که تفاوت عملکردشان در شکل‌ها مشخص باشد).

• در انتها و پس از پیاده‌سازی این الگوریتم توسط خودتان، یک بار هم از کلاس موجود در کتابخانه استفاده کنید و نتایج مدل خودتان و مدل آماده را مقایسه کنید.

## فاز دوم: RandomForest

در این فاز از شما میخواهیم تا الگوریتم RandomForest را پیاده سازی کنید و بر روی دیتاست ویژوال کرده و بررسی کنید. مجدداً فقط برای ساخت درخت تصمیم میتوانید از کتابخانه sklearn استفاده کنید.

- خروجی الگوریتم و ۵ درخت Randomized شده را بر روی دیتاست ویژوال کرده و بررسی کنید.
- در انتها و پس از پیاده سازی این الگوریتم توسط خودتان، یک بار هم از کلاس موجود در کتابخانه استفاده کنید و نتایج مدل خودتان و مدل آماده را مقایسه کنید.

## فاز سوم: AdaBoost

در این فاز از شما میخواهیم تا الگوریتم AdaBoost را بر روی دیتاست اجرا کنید و مدلی آموزش دهید. برای این الگوریتم میتوانید از کتابخانه sklearn و توابع آماده آن استفاده کنید.

- 8 مرحله از اجرای آن را بر روی دیتاست ترسیم کرده و بررسی کنید که در آن مرحله کلسیفایرهای ضعیف چه عملکردی داشته‌اند چه بخش‌هایی را به درستی کلسیفای کرده و این بخش‌ها در طول انجام الگوریتم و در مراحل که ترسیم کرده‌اید چگونه بهبود یافته و تغییر کرده‌اند. (بهتر است این 8 مرحله با فاصله از یکدیگر انتخاب شوند تا تفاوت بین آنها بهتر مشخص شود). کلسیفایرهای ضعیف این بخش را به دلخواه خودتان تعیین کنید.

## فاز چهارم: Stacked Learners

در این فاز از شما می‌خواهیم که همانطور که در درس یاد گرفته‌اید با ترکیب چند طبقه بند، یک طبقه بند stacked ایجاد کنید. انتخاب طبقه بندها به اختیار خودتان است ولی حداقل 4 طبقه بند متفاوت داشته باشید و برای تمام آنها hyperparameter tuning انجام دهید. پیاده سازی این فاز تماماً به عهده خودتان است (به جز ایجاد طبقه بندهای اولیه) و خودتان باید آنها را با هم ترکیب کنید.

**نمره اضافه:** در صورت overfit شدن هر یک از طبقه بندهای استفاده شده در این فاز، با توجه به روش ارائه شده در درس و کلاس حل تمرین، از این اتفاق جلوگیری کنید. در صورتی که هیچ یک از طبقه بندهای شما overfit نشد(که به احتمال زیاد به علت hyperparameter tuning هیچ کدام overfit نمی‌شوند) یک طبقه بند که overfit شود به مجموعه خود اضافه کنید و سپس از overfit جلوگیری کنید.

- بهتر است از طبقه بندهای آموزش دیده شده در فازهای قبلی استفاده نکنید و برای این فاز طبقه بندهای جدیدی روی داده‌ها آموزش دهید.

## نکات تکمیلی

- حتما موارد ذکر شده در اول صورت پروژه را رعایت کنید. در غیر این صورت از شما نمره کسر خواهد شد.
- علاوه بر سورس کد پروژه، فایل مستندات نیز باید آپلود شود.
- نام اعضای گروه در فایل مستندات ذکر شود و فقط یکی از اعضا پروژه را آپلود کند.
- هر گونه شباهت نامتعارف بین کد شما و کد سایر گروه ها تقلب محسوب می شود و نمره ای برای این پروژه دریافت نخواهید کرد.
- در صورت نوشتن داکيومنت تمیز (برای مثال با LATEX) نمره اضافه برای شما در نظر گرفته خواهد شد.
- فایل شامل سورس کد پروژه و مستندات را در قالب فایل zip و با نام شماره دانشجویی خود ذخیره و ارسال نمایید.
- در صورت داشتن هرگونه سوال می توانید با [Kourosh\\_Hsz](#) و [Yosef\\_Rezazadeh81](#) در ارتباط باشید یا در گروه درسی مطرح کنید.

موفق باشید؛  
تیم حل تمرین