



## پروژه دوم درس مبانی یادگیری ماشین

### مدل دسته‌بندی و خوشه‌بندی برای برگ‌ها

فایل `leaves.csv` حاوی اطلاعات برگ‌های ۳۶ گیاه مختلف است که ستون اول آن، دسته مربوطه و بقیه ستون‌ها ویژگی‌های برگ‌هاست. مدل‌های مختلف برای دسته‌بندی و خوشه‌بندی این مجموعه داده را بررسی کرده و یک دسته‌بند و خوشه‌بند خوب ارائه نمایید. همچنین ارتباط خوشه‌های نهایی را با دسته‌ها بررسی کنید. (مثلاً اگر خوشه‌بندی  $K$  خوشه‌ای داشته باشیم، بیشتر تصاویر کدام دسته‌ها در هر خوشه قرار می‌گیرد؟)

#### حالت امتیازی:

علاوه بر فایل `CSV`، مجموعه تصاویر هر برگ نیز در پوشه `leaves.rar` موجود است که در این حالت می‌توانید از روی تصاویر موجود، خودتان ویژگی‌های دیگری نیز استخراج نمایید و به دیتاست اولیه اضافه نمایید.

#### نکات:

- ارسال فایل متنی گزارش عملکرد و کد نهایی و همچنین مستندات کد، الگوریتم و مدل نهایی (به صورت متن یا ویدیو) الزامی است.
- حتماً در مستندات خود روند الگوریتم نهایی حاوی پیش‌پردازش‌ها (حذف داده پرت، تمیز کردن دیتا، مهندسی و تحلیل ویژگی‌ها، مدل نهایی، تعیین ابرپارامترهای مدل نهایی و نحوه ارزیابی مدل و مقایسه با مدل‌های دیگر، چالش‌ها و پیشنهادات دیگر) را به طور کامل شرح دهید.
- فایل‌های خود را حداکثر تا تاریخ ۶ تیرماه به ایمیل [drmohamadkiani@gmail.com](mailto:drmohamadkiani@gmail.com) ارسال نمایید.
- به ازای هر روز تاخیر در ارسال، کسر نمره تا ۲۰ درصد منظور می‌گردد.
- به گزارش‌های مشابه نمره ای تعلق نمی‌گردد.
- محدودیتی در انتخاب پکیج و زبان برای حل مسئله وجود ندارد. اما در مستندات باید شرح داده شود.
- ارسال پروژه تک نفره می‌باشد و ارسال پروژه به صورت دو نفره تنها در صورتی مجاز است که مستندات و تحلیل‌ها بسیار جامع و کامل باشد و همچنین شامل حالت امتیازی نیز باشد وگرنه نمره پروژه تقسیم می‌گردد.

خوشدودن باشد. [Drmkiani.ir](http://Drmkiani.ir)