

پروژه درس هوش محاسباتی:

دیتاستی که در اختیار شما قرار داده شده است LFW یا Labeled Faces in the Wild، دیتاست شاخته شده از صورت اشخاص معروف می باشد. برای کسب اطلاعات بیشتر در مورد این دیتاست می توانید به لینک زیر مراجعه کنید:

<http://vis-www.cs.umass.edu/lfw/>

در پوشه lfw تعدادی پوشه مشاهده میکنید که در هر پوشه تعدادی تصویر قرار گرفته و هر پوشه براساس اسم فردی که پوشه حاوی تصاویر او است نام گذاری شده است. بازنمایی این تصاویر در قالب تعدادی فایل تکست در اختیار شما قرار داده شده است. از مدل VGGFace (<http://www.robots.ox.ac.uk/~vgg/publications/2015/Parkhi15/parkhi15.pdf>) برای استخراج ویژگی هر یک از تصاویر استفاده شده است و این ویژگی ها در یک فایل متنی با نام یکسان با هر عکس در کنار آن ذخیره شده اند.

با استفاده از مدل های Ensemble Learning که در کلاس آموختید، یک کلسیفایر ترین کنید که با گرفتن ویژگی های دو عکس (از اطلاعات ذخیره شده در فایل txt به عنوان ویژگی های هر عکس استفاده کنید.) به عنوان ورودی، بگوید این دو عکس متعلق به یک نفر هستند یا دو فرد متفاوت.

برای این کار نیازمند دو دسته دیتا هستیم. تعدادی زوج های دوتایی عکس که هر دو متعلق به یک نفر باشند (کلاس مثبت) و تعدادی زوج های دوتایی عکس که متعلق به دو فرد مختلف باشند (کلاس منفی). به همین منظور دو فایل تکست در کنار دیتاست قرار گرفته است، به نام pairsDevTrain و pairsDevTest. در هر سطر از این فایل ها در ابتدا نام یک فرد نوشته شده و روبروی آن دو عدد. مثلا در pairsDevTest داریم:

```
500
Abdullah_Gul      13      14
Abdullah_Gul      13      16
Abdullatif_Sener      1        2
Adel_Al-Jubeir     1        3
Al_Pacino          1        2
Alan_Greenspan     1        5
Albert_Costa       2        6
Albert_Costa       4        6
```

سطر اول بیانگر این است که ۵۰۰ داده برای کلاس اول و ۵۰۰ داده برای کلاس دوم خواهیم داشت. سپس در خط بعد نام Abdullah_Gul به همراه دو عدد ۱۳ و ۱۴ آمده است. این یعنی اولین داده ترین را تصویر ۱۳ و ۱۴ قرار گرفته در پوشه Abdullah_Gul در نظر بگیرید و لیبل آن مثبت (عکس های متعلق به یک فرد). به همین ترتیب ۵۰۰ داده کلاس مثبت را پشت سر هم داریم. سپس از سطر ۵۰۱ ام فایل به شکل زیر است:

```
AJ_Lamas          1      Zach_Safrin      1
Aaron_Guiel       1      Reese_Witherspoon 3
Aaron_Tippin      1      Jose_Luis_Rodriguez_Zapatero 1
Abdul_Majeed_Shobokshi 1      Charles_Cope     1
Abdullah_Gul      16     Steve_Cox        1
Abid_Hamid_Mahmud_Al-Tikriti 1      Eli_Broad        1
Adam_Kennedy      1      Amelie_Mauresmo 19
Adel_Al-Jubeir    1      Elisabeth_Welch 1
Adrian_Murrell    1      Tommy_Franks     15
Adriana_Lima      1      Terrence_Trammell 1
Adriana_Benez_Navarro 1      Jennifer_Cannata 0
```

سطر اول این قسمت از دیتا بیانگر این است که داده اول کلاس منفی شامل عکس اول در پوشه AJ_Lamas و عکس اول در پوشه Zach_Safrin است که با توجه به این که به اشخاص متفاوت تعلق دارند لیبل این جفت داده منفی خواهد بود.

با توجه به این اطلاعات از ۲۲۰۰ داده قرار گرفته در فایل pairsDevTrain برای آموزش سیستم خود استفاده کنید و سپس نتیجه را روی pairsDevTest گزارش کنید.

شما می توانید برای پیاده سازی پروژه از متلب یا پایتون استفاده کنید. لطفا کد و گزارش خود را تا پایان روز ۱۳ بهمن در ویو آپلود کنید. استفاده از کتابخانه های آماده بلامانع می باشد. هر گونه مشابهت نامتعارف بین کد دانشجویان و یا شباهت با کد های موجود در اینترنت به منزله تقلب در نظر گرفته شده و با متخلفین برخورد خواهد شد.