

### پروژه شماره ۳

۱. یک دیتاست به دلفواه دانلود کنید و به تحلیل آن بپردازید (این تحلیل به شما در بخش پیش پردازش داده ها یا انتساب روش طبقه بندی مناسب کمک خواهد کرد)
۲. با توجه به تحلیل انجام شده در مرحله قبل پیش پردازش مورد نیاز روی داده ها را انجام دهید (به عنوان مثال بر اساس تحلیل انجام شده در می یابید که داده ها miss value دارند یا خیر. به عنوان مثالی دیگر با تعیین نوع داده ی هر ویژگی در می یابید که در مورد سلول های null موجود بر چه اساس عملیات جایگزینی را انجام دهید و سلول های null را مقداردهی کنید).
۳. با روشی به جز pca کاهش بعد را در داده ها انجام دهید. عملیات کاهش بعد را به صورت الگوریتمی پیاده سازی کنید.
۴. طبقه بندی داده ها را با روشی دلفواه از روش های طبقه بندی (روشی که در کلاس آموزش داده نشده) به صورت الگوریتمی انجام دهید.
۵. نمودار confusion matrix فروجی داده ها را رسم نمایید.

### پروژه شماره ۳

۱. یک دیتاست با استفاده از کتابخانه پانداس ایجاد کنید.
۲. دیتاست ایجاد شده را با استفاده از یکی از روش های فوشه بندی (روشی که در کلاس آموزش داده نشده) به صورت الگوریتمی فوشه بندی کنید.
۳. با دیتاست پروژه شماره ۲ نیز عملیات فوشه بندی را انجام دهید.

### توضیحات و نکات:

۱. هدف از طرح سوالات به چالش کشیدن شما در قابلیت مهم سرچ است. برای این منظور شما می توانید از:
  - a. سایت های آموزشی (مثل w3schools)
  - b. داکيومنتیشن توسعه دهنده های کتابخانه های پایتون و ... (تمام کتابخانه ها و زبان های برنامه نویسی و ... سایت دارند که آن قسمت docs قوی ترین آموزش ها را ارائه می دهد)
  - c. سایت های پرسش و پاسخ (مثل stackoverflow)
۲. تمویل هفته شانزدهم به صورت مضوری در کلاس درسی می باشد.
۳. منظور از الگوریتمی: همانطور که در کلاس مطرح شد تمام روش ها با یک تابع قابل پیاده سازی هستند. فوایشندم مراحل پروژه به صورت مرحله ای و الگوریتمی انجام دهید و از استفاده تنها یک تابع بپرهیزید.
۴. تسلط به پروژه و الگوریتم ها از الزامات تمویل پروژه است و نمره بر این اساس صورت می گیرد.
۵. پاسنگو سوالات عزیزان هستیم.