

## تمرین سوم\_ مسافرین کشتی تایتانیک

به عنوان تمرین سوم، از شما می‌خواهیم با استفاده از مجموعه داده مربوط به لیست مسافرین کشتی تایتانیک (۳ عدد فایل ضمیمه شده) با استفاده از آموخته‌هایتان تا این مرحله از دوره، EDA و بصری‌سازی را اجرا نمایید. در این تمرین علاوه بر مرور مطالب قبلی، از شما انتظار می‌رود با استفاده از الگوریتم‌های طبقه بندی زنده ماندن یا نماندن مسافرین بر اساس ورودی های موجود را پیش‌بینی نمایید.



قدم اول:

- Import نمودن کتابخانه های مورد نیاز (کتابخانه هایی نظیر Numpy و Pandas و ...)

قدم دوم:

- بارگذاری و خواندن دیتاست با استفاده از کتابخانه Pandas
- اکتشاف دیتاست و بررسی داده به صورت کامل (EDA)

قدم سوم:

- شناسایی داده‌های گم شده و Missing Value ها

قدم چهارم:

- ترسیم نمودارهای مختلف جهت درک بهتر از داده‌ها با استفاده از Seaborn و یا کتابخانه‌های دیگر

قدم پنجم:

- آموزش مدل با استفاده از روش های طبقه بندی آموخته شده: (با استفاده از فایل train.csv)
  - Logistic Regression
  - KNN Regressor
  - Decision Tree
  - Random Forest
  - SVM
  - XGBoost
  - ....

قدم ششم:

- ذخیره مدل‌های نهایی و آپلود آن‌ها در سایت مکتب‌خونه به همراه کد نهایی در قالب یک Jupyter Notebook

اختیاری) قدرت الگوریتم خود را بر روی داده‌های test.csv اعمال نموده و جواب را گزارش دهید.

نمونه خروجی نهایی در فایل sampleSubmission.csv موجود می‌باشد.