

پروژه پایانی دوره یادگیری ماشین (رگرسیون)

نام پروژه: پیشبینی تعداد رای‌های منفی فیلم‌های یوتیوب

چکیده پروژه: در این پروژه دانشجویان باید با استفاده از مجموعه داده‌گان فراهم شده و استفاده از مدل‌های رگرسیون خطی یا غیر خطی، تعداد رای‌های منفی ویدئو ها (ستون dislike_count) را پیشبینی کند.

مراحل پروژه:

- ☐ بارگذاری داده‌های آموزشی در پایتون و پیش پردازش داده‌ها
- ☐ برازش چندین مدل رگرسیون و مقایسه آن‌ها
- ☐ مصورسازی دقت پیشبینی مدل‌های رگرسیون طراحی شده.
- ☐ مقایسه مدل‌های رگرسیون بر حسب انواع متریک‌های عددی
- ☐ انتخاب بهترین مدل و پیشبینی تعداد رای‌های منفی در داده‌های تست
- ☐ (اختیاری) بارگذاری فایل نوت بوک نهایی در کگل

<https://www.kaggle.com/datasets/adampsu/youtube-dislikes-info>

* برای انجام این پروژه می‌توانید از هر مدل رگرسیون خطی یا غیرخطی بجز شبکه‌های عصبی استفاده کنید.
* پروژه‌ها باید به صورت یک فایل جویپتر نوت بوک تحویل داده شوند. پروژه‌های ارسال در هر قالب دیگری مورد بررسی قرار نخواهند گرفت.

بخش امتیازی:

مدل‌های رگرسیون همواره مقادیر پیوسته را می‌توانند پیشبینی کنند. در نتیجه تمامی اعداد حقیق در فضای پیشبینی مدل‌ها قرار می‌گیرد. اما تعداد رای‌های مثبت و منفی (like , dislike) یک عدد صحیح است. در صورتی که دانشوران از تکنیکی استفاده کنند که پیشبینی مدل به صورت عدد صحیح باشد ۲۵ درصد امتیاز بیشتر از پروژه دریافت می‌کنند.