

به نام خدا

در ابتدا عرض کنم بنده توضیحات در این فایل را در ویدیو ضبط شده نیز داده ام.

سوال دوم، مورد سوم

✓ چرا از بیز ساده استفاده نکنیم بهتر است؟

به سبب اینکه این رویکرد طراحی شده است برای مواقعی که جنس ورودی ما از جنس category هستند بسیار مناسب است. این رویکرد یک رویکرد با نظارت (supervised) است که باعث میشود بر اساس برچسب ها در حین یادگیری ارتباط میان ورودی و کلاس ها مشخص شود.

✓ چرا استفاده از k-mean مناسب نیست؟

الگوریتم k-mean برای داده هایی categorical زیاد مناسب نیست، حتی وقتی که ما این رسته ها را به داده های عددی تبدیل کردیم. الگوریتم k-mean یک رویکرد بدون نظارت است بدین معنی که کاری با برچسب داده ها ندارد و ما به صورت غیر مستقیم کلاس خوشه ها را برچسب های واقعی آنها بررسی کردیم. همچنین ما باید در انتخاب k بسیار دقت کنیم تا دچار خوشه های بی معنی نشویم و باید از قبل شناخت کافی به مجموعه داده ها داشته باشیم.