

ارزیابی و انتخاب ویژگی - تکلیف دوم

پس از استخراج ویژگی ها، نوبت به انتخاب ویژگی ها می‌رسد. برای انجام این مهم، شما باید یک روش کامل انتخاب ویژگی پیاده سازی کنید. برای این کار، از اجزای مختلف روش های انتخاب ویژگی بهره می‌گیریم:

- How well each feature acts on its own
- How different is each feature from the rest of features in general

در خصوص مورد اول، باید بررسی کنیم که یک ویژگی چقدر می‌تواند تمایز بین داده ها را نشان دهد. برای بررسی این مهم، از یک طبقه بند درخت تصمیم بهره می‌گیریم، و سعی می‌کنیم صرفاً با یک ویژگی، داده ها را طبقه بندی کنیم. دقت خروجی، یک معیار سنجش برای کیفیت ویژگی خواهد بود.

در خصوص متریک دوم، از معیار correlation برای بررسی تک ویژگی استفاده می‌کنیم (مقایسه دو به دو). ویژگی خوب، ویژگی ای هست که کمترین correlation را با بقیه داشته باشد (نمره اضافه : مقایسه با کل ویژگی ها - توضیح در کلاس حل تمرین).

و نهایتاً دو عدد بدست آمده باید با هم ترکیب شوند. برای ترکیب دو عدد، ابتدا آنها را نرمال کرده، و سپس از یک روش مشابه روش محاسبه f_1 استفاده کنید.

خروجی اصلی این بخش، ویژگی های انتخاب شده خواهد بود، همچنین در گزارش روال انتخاب ویژگی را توضیح دهید و سعی کنید توضیح دهید چه ویژگی هایی اطلاعات داشته اند و چه ویژگی هایی تکراری بوده‌اند.

بهبود دقت با بخش بندی¹ - تکلیف دوم

نهایتاً در بخش آخر، شما باید سعی کنید با استفاده از بخش بندی، دقت و سرعت مدل را بهبود ببخشید. در این مرحله، شما دیتا های خود را ابتدا با یک الگوریتم بخش بندی به بخش های مختلف می‌شکنید، و سپس سیستم قبلی را بر روی هر بخش مجزا آموزش می‌دهید، و در زمان تست، با توجه به اینکه داده ورودی در کدام بخش قرار می‌گیرد، از زیر سیستم مربوطه استفاده می‌کنید. نهایتاً یک سیستم یکتا برای تشخیص تشنج خواهید داشت.

¹ Clustering

ارزیابی با مسئله های مختلف - تکلیف دوم

همانطور که در جلسه حل تمرین توضیح داده شد، برای این دیتاست، مسئله های مختلفی وجود دارد. با توجه به توضیحات، ۳ حالت ۲ کلاس و یک حالت ۳ کلاس را انتخاب کنید، و روش خود را بر روی آنها ارزیابی کنید.