

تمرین Ensemble by Voting

محمد صالح علی اکبری

تاریخ تحویل: ۲ خرداد

سؤال ۱

انواع Voting را تشریح کرده و جمع‌بندی نهایی را ارائه دهید.

پاسخ

در حوزه یادگیری ماشین، یکی از روش‌هایی که برای بهبود عملکرد مدل‌ها استفاده می‌شود، Voting است؛ یعنی ترکیب پیش‌بینی چند مدل مختلف برای رسیدن به یک پاسخ نهایی. این روش به‌طور کلی به دو نوع اصلی تقسیم می‌شود:

- رأی‌گیری سخت (**Hard Voting**): در این روش، هر مدل صرفاً یک برچسب (کلاس) برای نمونه مورد نظر پیش‌بینی می‌کند و در نهایت، آن کلاسی که بیشترین رأی را آورده باشد به عنوان خروجی نهایی انتخاب می‌شود. مثل رأی‌گیری در انتخابات!

- رأی‌گیری نرم (**Soft Voting**): اینجا مدل‌ها به‌جای برچسب، احتمال تعلق نمونه به هر کلاس را اعلام می‌کنند. این احتمال‌ها با هم ترکیب می‌شوند و کلاسی که مجموع احتمالش بیشتر باشد، برنده است.

به‌جز این دو مورد، یک دسته دیگر از روش‌ها وجود دارد که به آن‌ها Non-trainable Voting می‌گویند. در این روش‌ها خبری از یادگیری نیست؛ فقط خروجی مدل‌ها با روش‌های ریاضی با هم ترکیب می‌شود. مثال‌هایی از این روش‌ها:

- میانگین‌گیری: همه خروجی‌ها را جمع می‌کنیم و بر تعداد مدل‌ها تقسیم می‌کنیم.

$$y_i = \frac{1}{L} \sum_{j=1}^L d_{ji}$$

- میانگین وزنی: هر مدل وزن خاص خودش را دارد؛ یعنی بعضی مدل‌ها مهم‌ترند.

$$y_i = \sum_j w_j d_{ji}, \quad w_j \geq 0, \quad \sum_j w_j = 1$$

- میانه: خروجی‌ها را مرتب می‌کنیم و مقدار وسط را انتخاب می‌کنیم.

$$y_i = \text{median}_j \{d_{ji}\}$$

- حداقل و حداکثر: به ترتیب کمترین یا بیشترین مقدار خروجی بین مدل‌ها را در نظر می‌گیریم.

$$y_i = \min_j \{d_{ji}\}, \quad y_i = \max_j \{d_{ji}\}$$

- ضرب: همه خروجی‌ها را در هم ضرب می‌کنیم؛ اگر یکی از مدل‌ها خروجی پایینی بدهد، روی نتیجه نهایی تأثیر زیادی خواهد داشت.

$$y_i = \prod_j d_{ji}$$

در این فرمول‌ها، d_{ji} یعنی پیش‌بینی مدل j برای کلاس i ، و w_j وزن اختصاص داده‌شده به مدل j است.

سؤال ۲

کاهش خطر انتخاب یک مدل ناکافی را تشریح کنید.

پاسخ

وقتی فقط از یک مدل استفاده می‌کنیم، ممکن است آن مدل روی همه داده‌ها عملکرد خوبی نداشته باشد. اما استفاده از ترکیب چند مدل (یعنی Ensemble Learning) باعث می‌شود تا ضعف‌های مدل‌ها تا حد زیادی پوشش داده شود. چند دلیل برای این موضوع:

- کاهش واریانس: اگر یک مدل خیلی نوسان دارد و در برابر داده‌های مختلف رفتار متفاوتی نشان می‌دهد، ترکیب آن با مدل‌های دیگر باعث می‌شود خروجی پایدارتر شود.
- کاهش بایاس: بعضی مدل‌ها ساده هستند و نمی‌توانند الگوهای پیچیده را یاد بگیرند. ترکیب آن‌ها با مدل‌های دیگر می‌تواند این مشکل را کاهش دهد.
- دقت بالاتر: معمولاً ترکیب چند مدل بهتر از عملکرد بهترین مدل منفرد جواب می‌دهد.

- پوشش بیشتر داده‌ها: هر مدل ممکن است روی بخش خاصی از داده‌ها خوب عمل کند. ترکیب آن‌ها کمک می‌کند تا همه‌ی جوانب داده‌ها بهتر پوشش داده شود.

یکی از روش‌های پیشرفته در این زمینه Bayesian Model Combination است. این روش با در نظر گرفتن احتمال‌ها و عدم قطعیت‌ها، وزن‌های مناسبی به مدل‌ها می‌دهد تا نتیجه ترکیبی بهینه‌تری حاصل شود. به این ترتیب، خطر انتخاب اشتباه مدل به حداقل می‌رسد.