

سوال 1) تکنیک MRMR در انتخاب ویژگی را تعریف کنید

MRMR مخفف Minimum Redundancy Maximum Relevance است که به معنای حداقل افزونگی و حداکثر ارتباط است. این روش به دنبال یافتن ویژگی‌هایی است که بیشترین ارتباط را با متغیر هدف داشته باشند و در عین حال کمترین همبستگی را با سایر ویژگی‌های انتخاب شده داشته باشند. به عبارت دیگر، MRMR سعی می‌کند تعادلی بین ارتباط ویژگی‌ها با متغیر هدف و استقلال آن‌ها ایجاد کند.

مفهوم اصلی MRMR

- ارتباط: ارتباط یک ویژگی با متغیر هدف معمولاً با استفاده از معیارهایی مانند اطلاعات متقابل اندازه‌گیری می‌شود. اطلاعات متقابل نشان می‌دهد که دانستن مقدار یک ویژگی تا چه اندازه عدم قطعیت در مورد مقدار متغیر هدف را کاهش می‌دهد.
- افزونگی: افزونگی بین دو ویژگی با استفاده از اطلاعات متقابل اندازه‌گیری می‌شود. اگر دو ویژگی اطلاعات متقابل بالایی داشته باشند، به این معنی است که آن‌ها همبستگی بالایی با هم دارند و به عبارت دیگر، اطلاعات تکراری را ارائه می‌دهند.

الگوریتم MRMR

1. ابتدا: یک مجموعه خالی از ویژگی‌های انتخاب شده در نظر بگیرید.
2. انتخاب بهترین ویژگی: ویژگی‌ای را که بیشترین ارتباط را با متغیر هدف دارد، انتخاب کرده و آن را به مجموعه ویژگی‌های انتخاب شده اضافه کنید.
3. تکرار: تا زمانی که تعداد مورد نظر از ویژگی‌ها انتخاب شود، مراحل زیر را تکرار کنید:
  - برای هر ویژگی باقی‌مانده، ارتباط آن با متغیر هدف و افزونگی آن با تمام ویژگی‌های انتخاب شده را محاسبه کنید.
  - ویژگی‌ای را که مقدار MRMR آن بیشترین باشد، انتخاب کنید. مقدار MRMR برابر است با تفاضل بین ارتباط با متغیر هدف و میانگین افزونگی با ویژگی‌های انتخاب شده.
  - ویژگی انتخاب شده را به مجموعه ویژگی‌های انتخاب شده اضافه کنید.

مزایای MRMR

- کارایی: الگوریتم MRMR نسبتاً ساده و سریع است.
- موثر: این روش در بسیاری از کاربردها نشان داده است که می‌تواند ویژگی‌های مفید و غیرتکراری را انتخاب کند.
- انعطاف‌پذیری: می‌توان از معیارهای مختلفی برای اندازه‌گیری ارتباط و افزونگی استفاده کرد.

## محدودیت‌های MRMR

- الگوریتم حریصانه MRMR: یک الگوریتم حریصانه است و ممکن است همیشه بهترین زیرمجموعه از ویژگی‌ها را پیدا نکند.
- حساسیت به نویز MRMR: می‌تواند به نویز در داده‌ها حساس باشد.
- فرض استقلال شرطی MRMR: فرض می‌کند که ویژگی‌ها به صورت شرطی مستقل از یکدیگر هستند که ممکن است همیشه صادق نباشد.

## کاربردهای MRMR

- کاهش ابعاد داده: با حذف ویژگی‌های غیرضروری، ابعاد داده کاهش می‌یابد که منجر به بهبود کارایی الگوریتم‌های یادگیری ماشین می‌شود.
- بهبود دقت مدل: با انتخاب ویژگی‌های مرتبط، دقت مدل‌های یادگیری ماشین بهبود می‌یابد.
- تفسیر بهتر مدل: با انتخاب تعداد کمی از ویژگی‌های مهم، تفسیر مدل‌های یادگیری ماشین آسان‌تر می‌شود.

## سوال (2) محاسبه ی MSE بر حسب واریانس و بایاس

خطای میانگین مربعات را می‌توان به صورت زیر به بایاس و واریانس تجزیه کرد:

$$MSE = Bias^2 + Variance + Noise$$

## اثبات

فرض کنید:

- $f(x)$ : تابع واقعی
- $\hat{f}(x)$ : تابع تخمین زده شده توسط مدل
- $y$ : مقدار واقعی
- $\epsilon$ : نویز تصادفی

خطای یک نمونه به صورت زیر محاسبه می‌شود:

$$Error = (f(x) - y)^2$$

با جایگزینی  $y = f(x) + \epsilon$  داریم:

$$Error = (f(x) - f(x) - \epsilon)^2$$

با باز کردن عبارت بالا و گرفتن امید ریاضی از هر دو طرف:

$$E[Error] = E[(f(x) - \hat{f}(x))^2] + E[\varepsilon^2] + 2E[(f(x) - \hat{f}(x))\varepsilon]$$

با فرض استقلال نویز از پیش‌بینی مدل، جمله‌ی آخر صفر می‌شود. بنابراین:

$$MSE = E[(f(x) - \hat{f}(x))^2] + E[\varepsilon^2]$$

جمله‌ی اول را می‌توان به بایاس و واریانس تجزیه کرد:

$$MSE = Bias^2 + Variance + Noise$$

که در آن:

•  $Bias^2 = (E[\hat{f}(x)] - f(x))^2$  نشان‌دهنده‌ی تفاوت بین مقدار پیش‌بینی‌شده‌ی متوسط مدل و مقدار واقعی است.

•  $Variance = E[(\hat{f}(x) - E[\hat{f}(x)])^2]$  نشان‌دهنده‌ی تغییرات پیش‌بینی‌های مدل حول مقدار متوسط آن است.

•  $Noise = E[\varepsilon^2]$  نشان‌دهنده‌ی نویز ذاتی در داده‌ها است که نمی‌توان آن را با مدل حذف کرد.