Viés–Variância: Estrutura Estatística do Erro em aprendizado de máquina.

Nos textos anteriores formalizamos o erro como risco empírico e risco esperado, além da geometria da função de perda L( ) [1].

2026-02-19

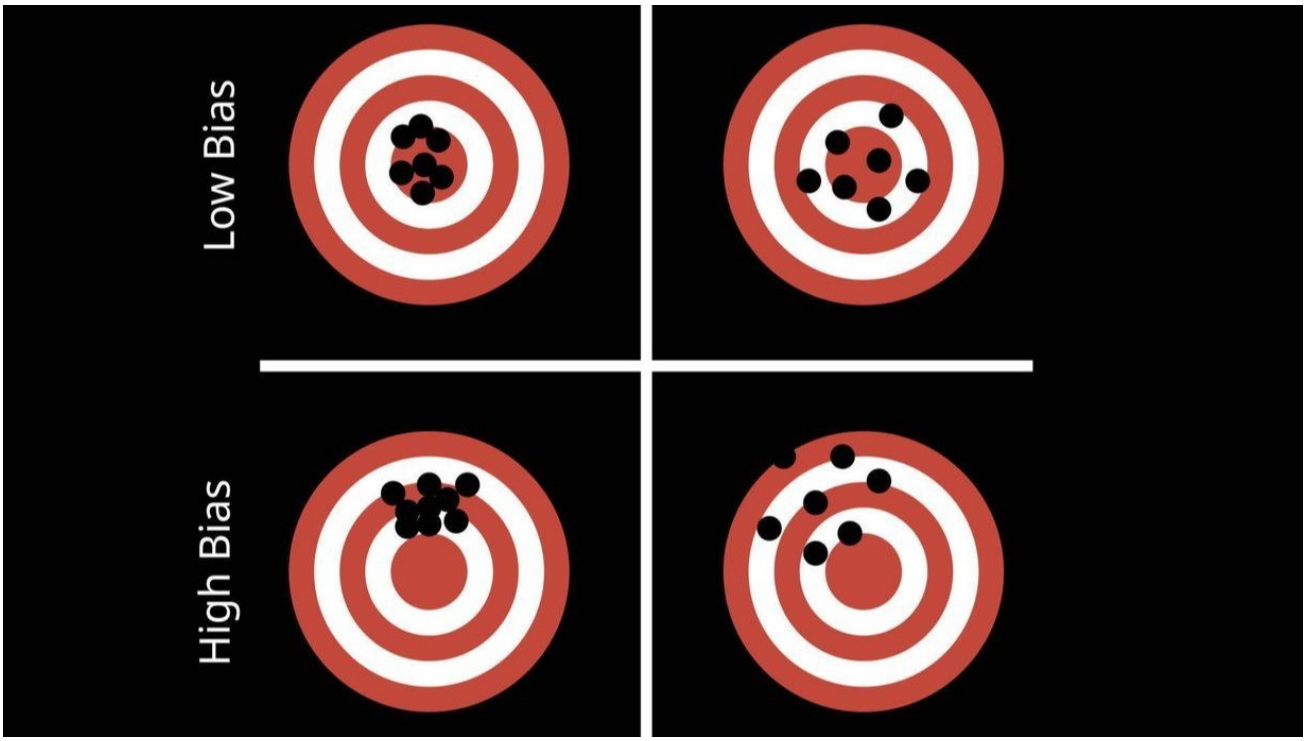


Figure 1: Capa do artigo

Nos textos anteriores formalizamos o erro como risco empírico e risco esperado, além da geometria da função de perda L( ) [1].

Aprofundamos agora a estrutura estatística desse erro por meio da decomposição viés–variância. Ela organiza formalmente o problema da generalização.

A Decomposição

Para um estimador f̂(x) que aproxima a função real f(x), o erro quadrático esperado pode ser escrito como:

$$\mathbb{E}\left[(y - \hat{f}(x))^2\right] = \text{Bias}^2(x) + \text{Var}(x) + \sigma^2$$

onde:

© Artemisia Weyl 2024–2026 — Todos os direitos reservados.

$$\text{Bias}(x) = \mathbb{E}[\hat{f}(x)] - f(x)$$

$$\text{Var}(x) = \mathbb{E}\Big[(\hat{f}(x) - \mathbb{E}[\hat{f}(x)])^2\Big]$$

^{2} representa o ruído intrínseco do processo gerador [1], [2]. Essa identidade separa o erro em componentes com interpretação própria.

Viés

O viés , aqui tambem chamado de BIAS mede o desvio sistemático do estimador em relação à função real.Ele surge quando o modelo impõe uma estrutura que não corresponde ao processo gerador.

Em aplicações reais, o viés está associado a:

- Variáveis relevantes não observadas
- Amostragem não representativa
- Dados incompletos

O resultado é uma aproximação deslocada da estrutura real.

Variância

A variância mede a sensibilidade do modelo à amostra utilizada no treinamento. Formalmente:

$$\text{Var}(x) = \mathbb{E}\Big[(\hat{f}(x) - \mathbb{E}[\hat{f}(x)])^2\Big]$$

Modelos com maior flexibilidade apresentam maior variabilidade entre diferentes conjuntos de treino [1]. No plano dos dados, variância elevada está associada a:

- Rótulos ruidosos
- Inconsistência entre fontes
- Outliers não tratados

Esse termo depende tanto da arquitetura quanto da qualidade da base.

$$Otermo\sigma^2$$

O componente ^{2} corresponde ao ruído de Bayes. Ele representa a variabilidade inerente ao processo, mesmo sob modelo ideal. Esse limite depende da relação sinal–ruído e da precisão da observação [5]. Melhorias na captura e na governança deslocam esse limite teórico.

Regularização

A regularização introduzida anteriormente,

$$L_{\text{reg}}(\theta) = L(\theta) + \lambda\|\theta\|^2$$

atua principalmente sobre a variância [2]. O parâmetro controla o equilíbrio entre estabilidade e flexibilidade. Esse ajuste modifica diretamente a decomposição do erro.

© Artemisia Weyl 2024–2026 — Todos os direitos reservados.

Regimes de Alta Dimensão

Em modelos superparametrizados observa-se o fenômeno de double descent [7]. Após o regime clássico de aumento de variância, o erro pode voltar a diminuir. Esse comportamento está relacionado à dinâmica de otimização e à estrutura espectral do problema.

A análise nesse regime exige considerar capacidade, ruído e dinâmica do treinamento.

Qualidade de Dados e Estrutura do Erro

Falhas na coleta e rotulagem propagam-se ao longo do pipeline. Oconceito de Data Cascades descreve como problemas upstream ampliam viés e variância downstream [8].

A decomposição viés–variância funciona, nesse contexto, como instrumento diagnóstico.Ela permite localizar a origem estrutural do erro.

CONCLUSÕES

A decomposição viés–variância organiza o erro esperado em três componentes:

1. Estrutura imposta pelo modelo
2. Variabilidade induzida pela amostra
3. Ruído inerente ao processo de geração dos dados

A arquitetura influencia o viés.

A composição da base influencia a variância.

A precisão da observação determina o limite associado a ^{2}.

Ao longo desta série, o erro foi tratado como objeto estatístico, geométrico e estrutural. decomposição viés–variância consolida essas dimensões em um mesmo quadro formal. Ela articula modelagem, otimização e engenharia de dados dentro de uma mesma estrutura analítica.

Referências

[1] T. M. Mitchell, *Machine Learning*, 1997.

[2] S. Boyd and L. Vandenberghe, *Convex Optimization*, 2004.

[3] C. M. Bishop, *Pattern Recognition and Machine Learning*, 2006.

[4] N. Sambasivan et al., "Data Cascades," CHI, 2021.

[5] S. Teh and K. Kempa-Liehr, IEEE Trans. Industrial Informatics, 2020.

[6] D. Anderson and M. Bjarnadottir, PLoS ONE, 2024.

[7] B. Nakkiran et al., "Deep Double Descent," 2021.

[8] A. Ng, "Data-Centric AI," 2021.

Postagem original no LinkedIn: Viés–Variância: Estrutura Estatística do Erro em aprendizado de máquina.

© Artemisia Weyl 2024–2026 — Todos os direitos reservados.