In this article, Professor Leo Breiman discusses the difference between two approaches of statistical modeling: the data modeling culture and the algorithmic modeling culture. The data modeling approach makes assumptions about data distribution and the algorithmic modeling approach is often non-parametric. Though the data modeling culture is more popular, Professor Leo Breiman argues through a few real-world examples that algorithmic models can provide better predictive power and more reliable interpretation.

The popularity in data science research is like reincarnation: over 20 years ago Professor Leo Breiman advocated that the algorithmic modeling approach should receive more attention but in 2020s, neural networks, which is one kind of algorithmic modeling approach, are being discussed in almost every application domain. Now under the popularity of the algorithmic modeling culture, people attempt to interpret complicated neural networks using simple logistic/linear regression models which are data modeling approaches [1,2]. It feels likes the data science community has been swinging between the two cultures and may find a balance someday in the future.

Reference:
[1] Rigotti M, Miksovic C, Giurgiu I, Gschwind T, Scotton P. Attention-based interpretability with concept transformers. InInternational Conference on Learning Representations 2021 Oct 6.
[2] Koh PW, Nguyen T, Tang YS, Mussmann S, Pierson E, Kim B, Liang P. Concept bottleneck models. InInternational conference on machine learning 2020 Nov 21 (pp. 5338-5348). PMLR.