Discussion of "Loss based prior for BART models"
Prof. Fabrizio Leisen

Luciana Dalla Valle

University of Turin (Italy)

luciana.dallavalle@unito.it

O'Bayes 2025

10 June 2025

Important Contributions of the Paper

Many congratulations to the authors (F. Serafini, C. Villa, F. Leisen and K. Wilson) for the important contributions of the paper:

- Novel Loss-Based (LB) prior for the tree topology of BART and CART models.
- The LB prior is a regularization prior, encouraging parsimonious and sparse trees, hence helping to avoid overfitting
- The LB prior can be calibrated in a principled way by maximizing the expected loss.
- With simulated and real data, the LB prior is advantageous over the classical prior

Questions for Discussion

Question n. 1

When introducing the LB prior for BART, the loss in complexity is defined as

$$Loss_C(T) = -\omega n_L(T) - \gamma\Delta(T)$$

where $\omega \geq 0$, $\gamma \in \mathbb{R}$, $n_L(T)$ is the number of terminal nodes of T and $\Delta(T)$ is the difference between the the number of left an right terminal nodes.
Why did you use $\Delta(T)$ and not the tree depth?

Questions for Discussion

Question n. 2

The number of trees m is a tuning parameter of the model. Could the LB prior be applied to the number of trees?

Question n. 3

In order to define the LB prior objectively, the expected loss was maximized. How does the expected loss method work in detail?

Questions for Discussion

Question n. 4

Villa and Lee (2020)[1] used the loss-based approach to design objective priors for variable selection in linear regression.
Could your ideas be applied to variable section with trees?

Question n. 5

Are computational times of the LB prior comparable to the classical prior?

[1] C. Villa and J. E. Lee. (2020) A Loss-Based Prior for Variable Selection in Linear Regression Methods. *Bayesian Analysis*, 15(2):533–558.