The story so far…

1. What is anomaly detection?

2. Tree-based methods?

3. Isolation Forests (and their scoring region **issue**)

4. Extended Isolation Forests (and their scoring region **solution**)

5. IF vs EIF (detection performance)

6. IF vs EIF (visualising scoring regions)

… and now we’re here!

Conclusion

EIF performance?

A closer look at IF vs EIF, along with the more traditional methods, should be trialled across a large variety of datasets. This thyroid dataset *preferring* splits across one dimension is likely one reason to it *not living up to expectations*.

And other method too !

Why you're at it why not try fit some of the distance & density based methods too. The more information you have the better; just always keep interpretation in mind

Use IF and EIF !

If either method detects an observation as being an anomaly then you can be more confident it's data worth your attention.