

Value Conflicts

Fundamentals of Artificial Intelligence Programme

Prof. Dr. Ibo van de Poel



Four types of approaches



Calculative approaches

- “Only the best”



Satisficing

- “Good enough”



Respecification

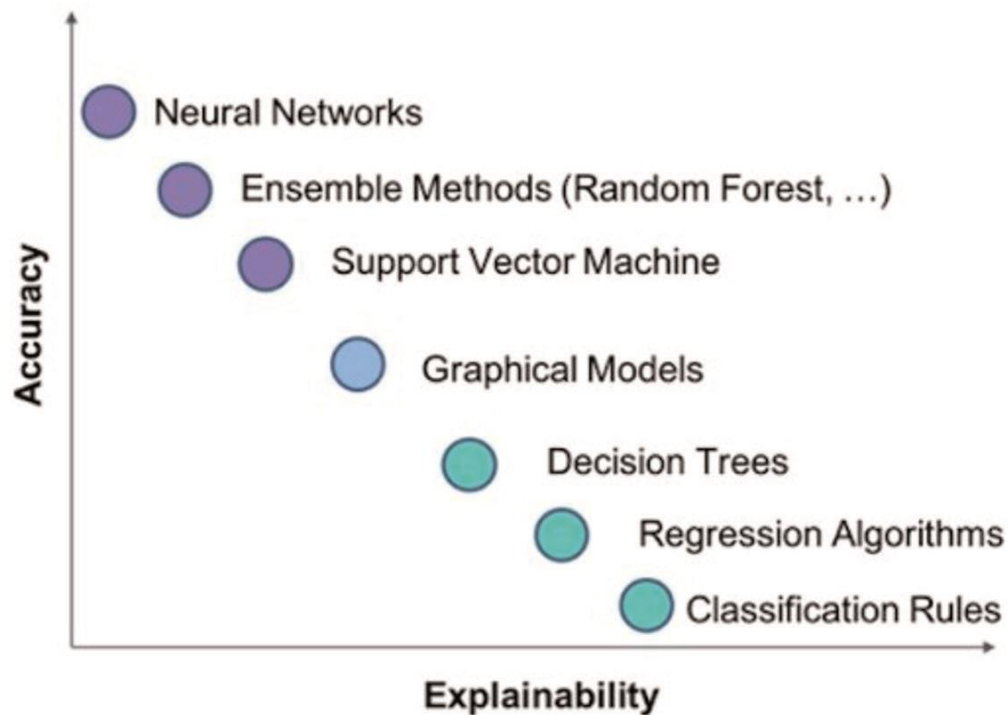
- “Think again”



Innovation

- “Try harder”

Trade-off between accuracy and explainability



Calculative Approach

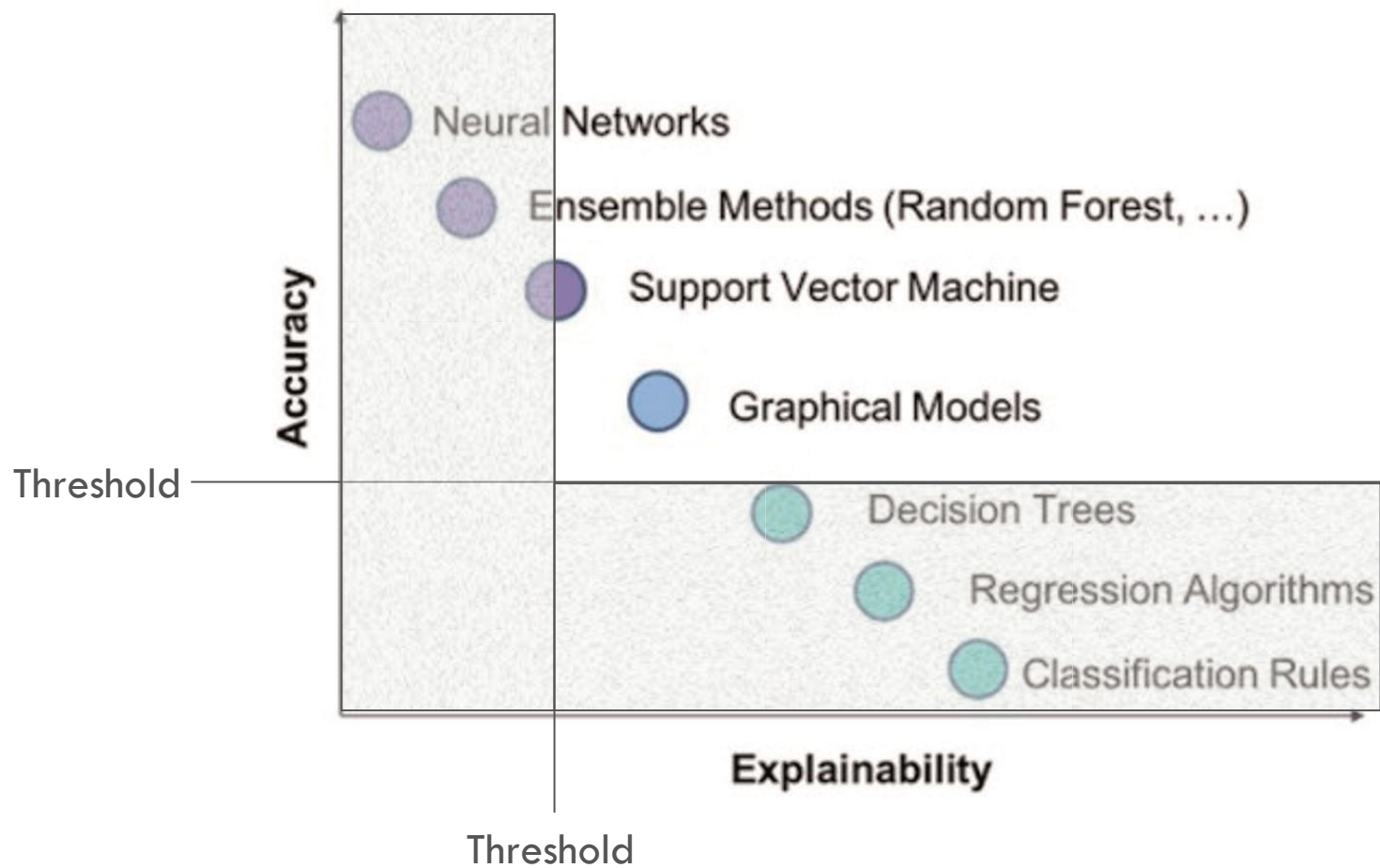
	Accuracy	Explainability	Overall score
Neural network	10	1	11
Random forest	9	3	12
Graphical model	7	7	14
Regression	5	8	13
Classification	2	10	12

Satisficing

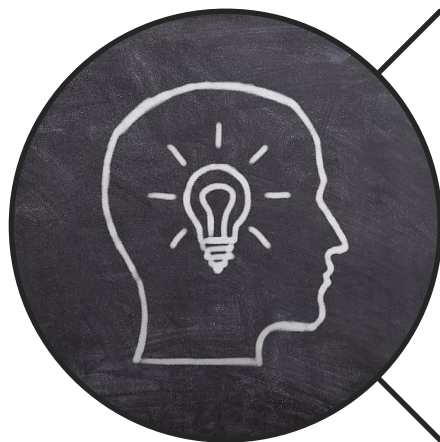


Satisficing

- “Good enough”



Respecification



Respecification

- “Think again”

Fairness versus Privacy



Fairness versus Privacy

Privacy

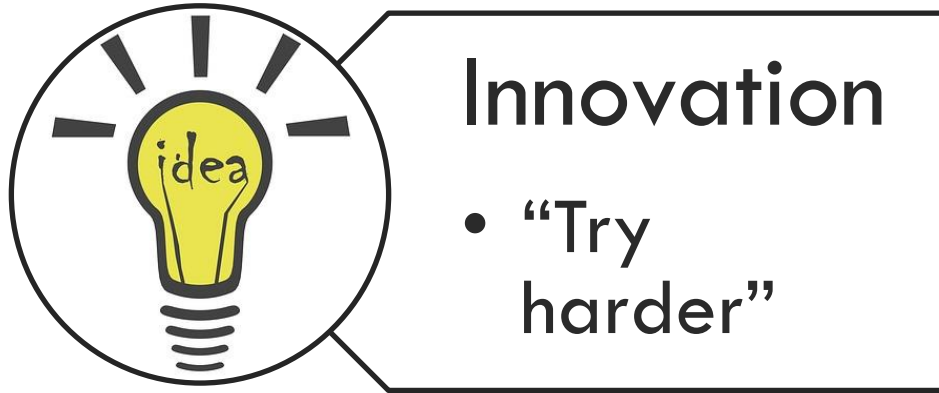


Confidentiality



Informed consent

Innovation



Summary



Calculative approaches

- “Only the best”



Satisficing

- “Good enough”



Respecification

- “Think again”



Innovation

- “Try harder”