



Ética no data science

Leonardo Alves (las3)



Is it ethical to avoid error analysis?

Eva García-Martín
Niklas Lavesson

<https://arxiv.org/pdf/1706.10237.pdf>



Falsificação

- Modificar ou omitir resultados e dados
- Cherry picking
- Análise superficial
- Não intencional



Discriminação

- Direta x indireta
- Características não protegidas
- Justificativas por trás do comportamento



No deep learning

- Sigilo de mercado
- Interpretação de semântica



A survey on measuring indirect discrimination in machine learning

Aalto Indre Zliobaite

<https://arxiv.org/pdf/1511.00148.pdf>



Aprendizagem de máquina não discriminatório

- Descoberta
- Prevenção



Testes estatísticos

- Teste de inclinação de regressão
- Por diferença de teste de médias
- Diferença em proporções para dois grupos ou vários grupos.



Medidas absolutas

- Diferença média
- Diferença normalizada
- Área sob a curva taxa de impacto
- Relação elift e relação de probabilidades



Métricas condicionais

- Diferença inexplicável
- Medida de propensão
- Relação de Belift



Métricas estruturais

- Testes de situação
- Medida de consistência.



Resultados

- Métricas baseadas em relação
- Métricas baseadas em diferenças
- Uso de uma única métrica



Perguntas?