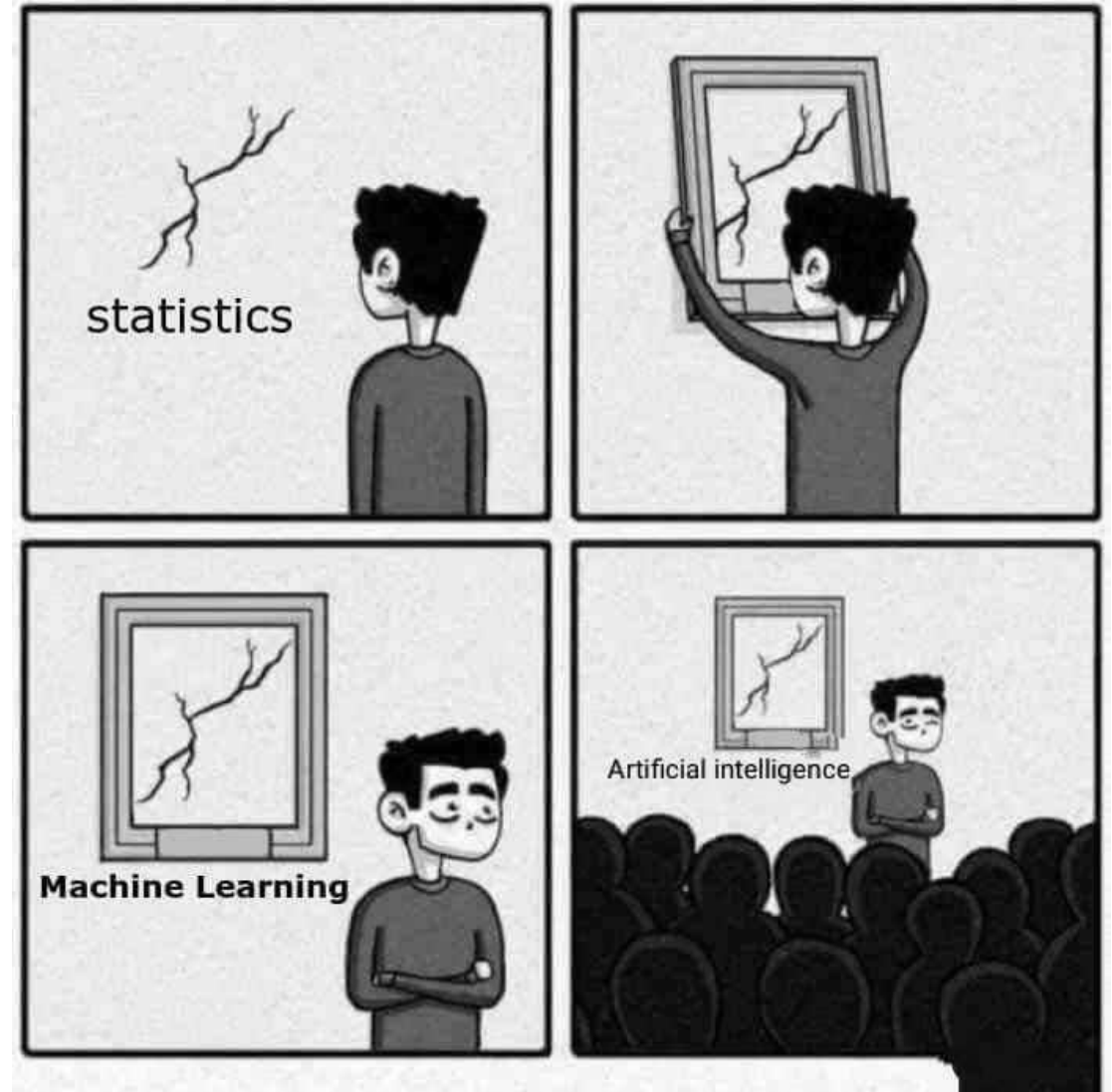


What is AI?

- Depends who you ask.
- Includes
 - Programmed AI algorithms (rule-based learning)
 - Machine Learning
 - (Statistical) techniques to "learn" some task with data, without being explicitly programmed
 - Machine "mimicking" human thinking
 - Autonomous system
 - Futuristic Robots
 - (A room full of interns)



What is Machine Learning?

Machine learning uses statistical techniques to give computers the ability to "learn" with data, without being explicitly programmed.



- Know something about the task
- Explicitly tell the machine how to complete it
- Test the model on something you know
- Deploy your program!



- Data Mining: Get data that you know something about
- Train the Model: "Teach" the machine about that data
- Test the model on something you know
- Deploy the program!

Potential Issues with AI Systems

- Lack of transparency
- ML often finds correlation, not causation
- Bias
- Base Rate Fallacy
- Privacy (e.g. membership attacks)
- Adversarial ML threats
- Result in antisocial and negative environmental outcomes
- Have adverse side effects
- Are built to only benefit a subset of users
- Externalize risks
- Produce errors due to distributional shift
- Result in systems that exploit states that fulfill the objective function, but do not complete the intended task
- Distribute errors unfairly
- ...