



# PROBLEM TRANSLATION

Business Analytics class

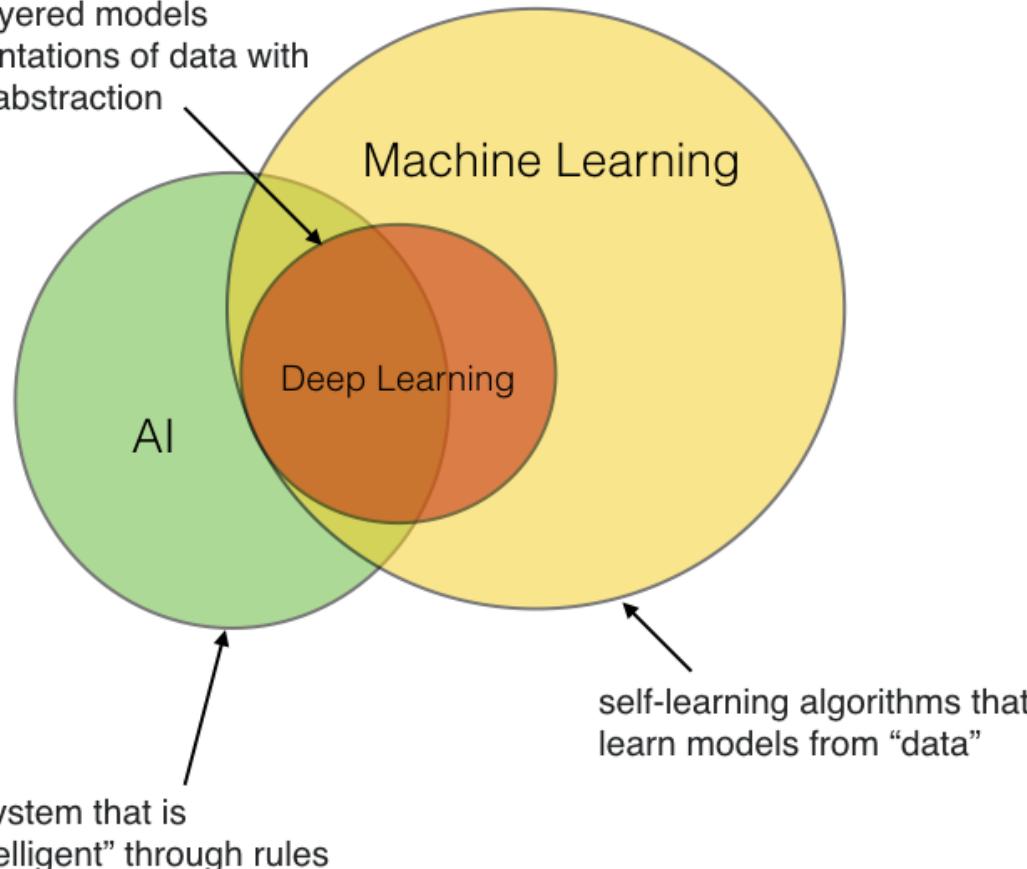
**What is data science?  
Artificial intelligence?  
Machine learning?  
Deep learning?  
Internet-of-things?  
Blockchain?**

**WHAT IT IS AND WHAT IT IS NOT**

# The Definitions

Things that can only learn from the past

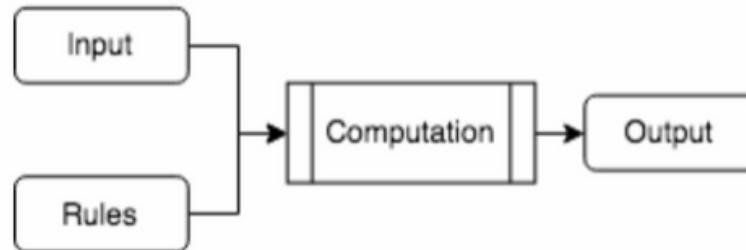
particular, multi-layered models  
that learn representations of data with  
multiple levels of abstraction



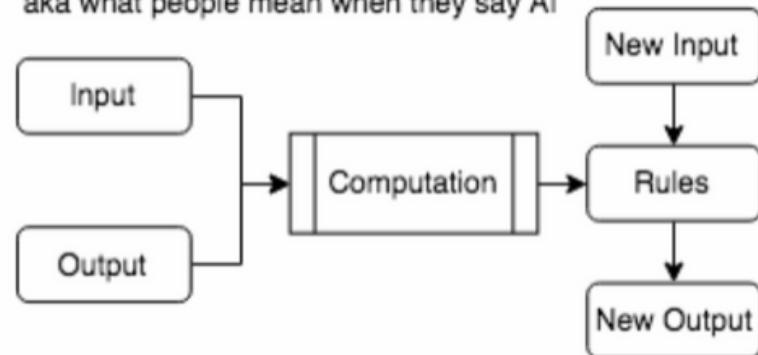
source: Sebastian Raschka. *Machine Learning FAQ - How are Artificial Intelligence and Machine Learning related?*. <https://sebastianraschka.com/faq/docs/ai-and-ml.html>

# What People Mean When They Say AI

Rule-based Programming



Differentiable Programming aka Machine Learning  
aka what people mean when they say AI



Which of these things belong in which categories, or none at all?

- Recommendation systems
- Search engines
- Sending images through chats
- Fraud detection
- Chatbots
- Snapchat filters
- Augmented reality
- Robot stock trading
- Ads pricing
- Health-risk evaluation
- Predictive maintenance
- Spell correction
- UX/UI design
- Social media notifications

---

## GROUP THINGS TOGETHER

Customer segmentation  
Product segmentation  
Product categorisation  
Fraud detection

# The Big Three

---

## PREDICT A NUMBER

Demand forecasting  
Customer scoring  
Dynamic pricing  
Sentiment analysis

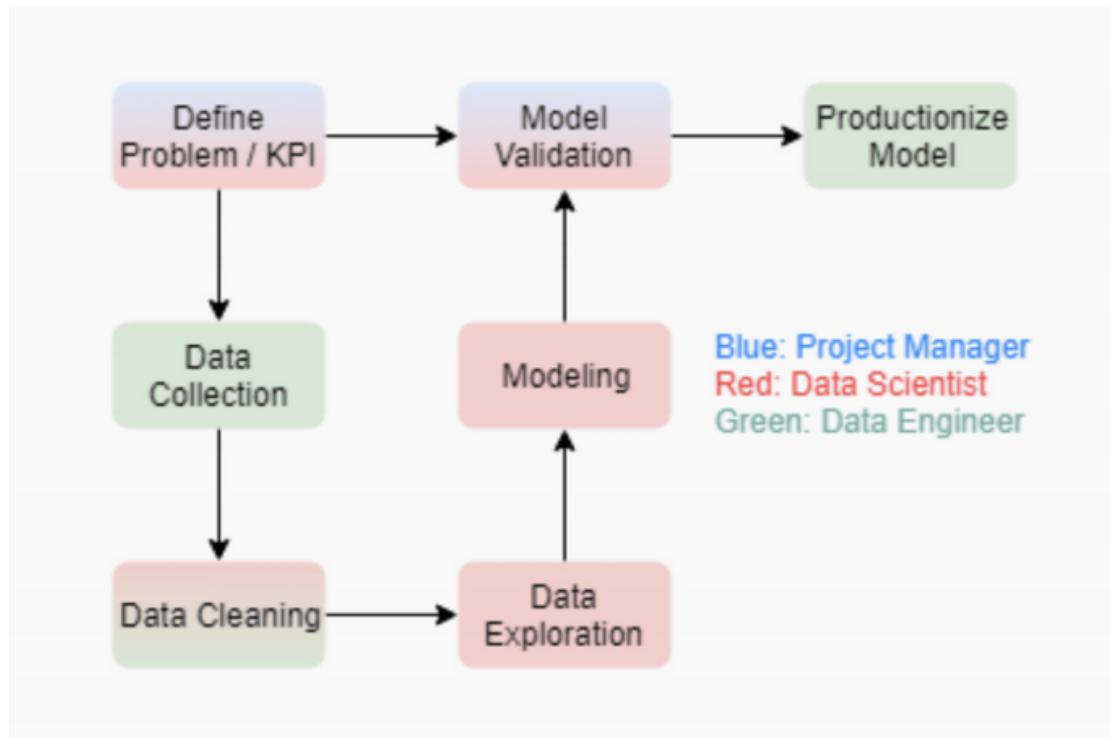
---

## NOT REALLY "AI"

A/B testing  
Trendspotting  
Market basket analysis  
Portfolio optimization

# Data Product Life Cycle

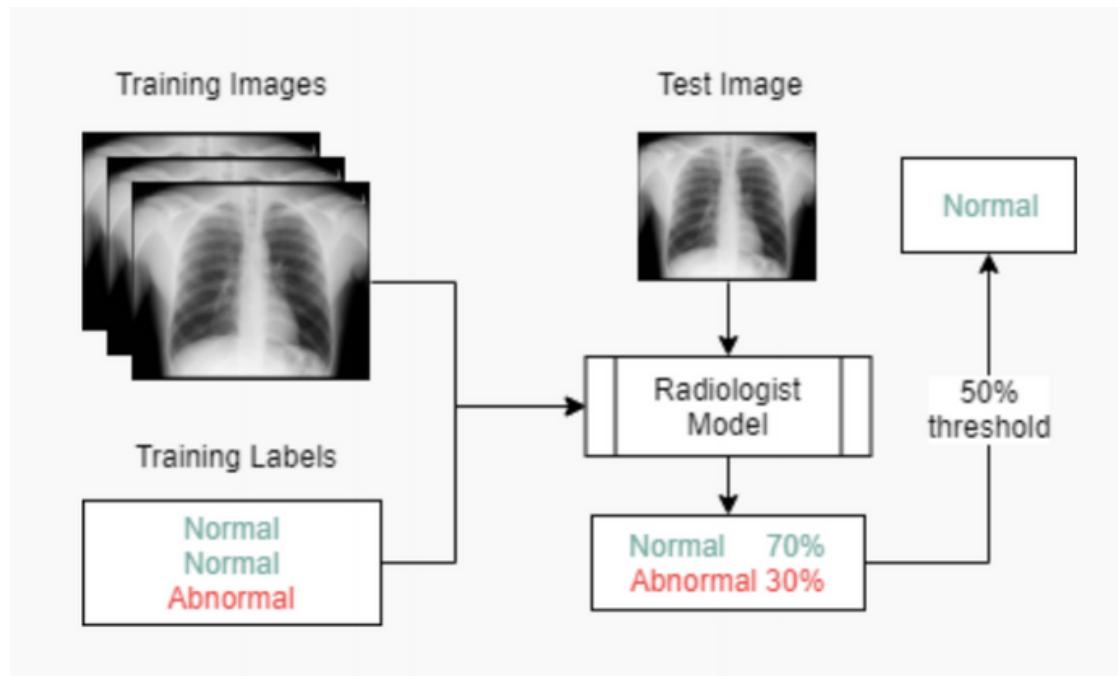
Who do you think is the most important party?



# What is a good performance?

Example: abnormality screening in chest x-rays

Is 97% accuracy good enough?



# How 97% accuracy will land you in prison

Example: abnormality screening in chest x-rays

N=100	True Normal	True Abnormal
Predicted Normal	97	3
Predicted Abnormal	0	0

N=100	True Normal	True Abnormal
Predicted Normal	80	0
Predicted Abnormal	17	3

---

## PREDICTED POSITIVE AT A GIVEN RECALL

How much effort can be saved  
given the same level of recall

---

## MEAN ABSOLUTE PERCENTAGE ERROR

What is the error in % that my  
model will make on average?

# Some useful realistic metrics

---

## ONLINE METRICS

Conduct A/B tests to confirm if  
your offline metrics reflect reality

# Why I have doubts about AutoML

## Aside from the fact that it is taking over my job

---

### COMPETITIVE ADVANTAGE

Can you really call yourself an AI company if you hire a freelancer to do AI for you?

---

### TROUBLE SHOOTING

What if something goes wrong and \$mn is on the line? AI usually works, until it does not

---

### DATA PROTECTION

How much trust do you have in Big Infra to not share your data and models with competitors

# Bottom line

