University of Essex

What is Machine Learning?

Akitaka Matsuo
Essex IADS

Machine Learning

University of Essex

What is machine learning?

- Learning from the data through algorithms/models

Types of machine learning

- **Supervised**: “building a statistical model for predicting, or estimating, an **output** based on one or more **inputs**”
  - **output**
    - exists in the data
    - similar to what we know as Dependent Variable
- **Unsupervised**: No output in the data. Models to learn the patterns in the data. Outputs might be generated by the model

Where to get output in supervised learning

University of Essex

- Maybe it already a part of the dataset
- It could be expensive to get the output data
  - Benoit et al (2016)
  - Anastasopulos and Bertelli (2019)
  - Manifesto data
    - Subamanian et al (2018)
- Who to code the data
  - Using cloud workers?

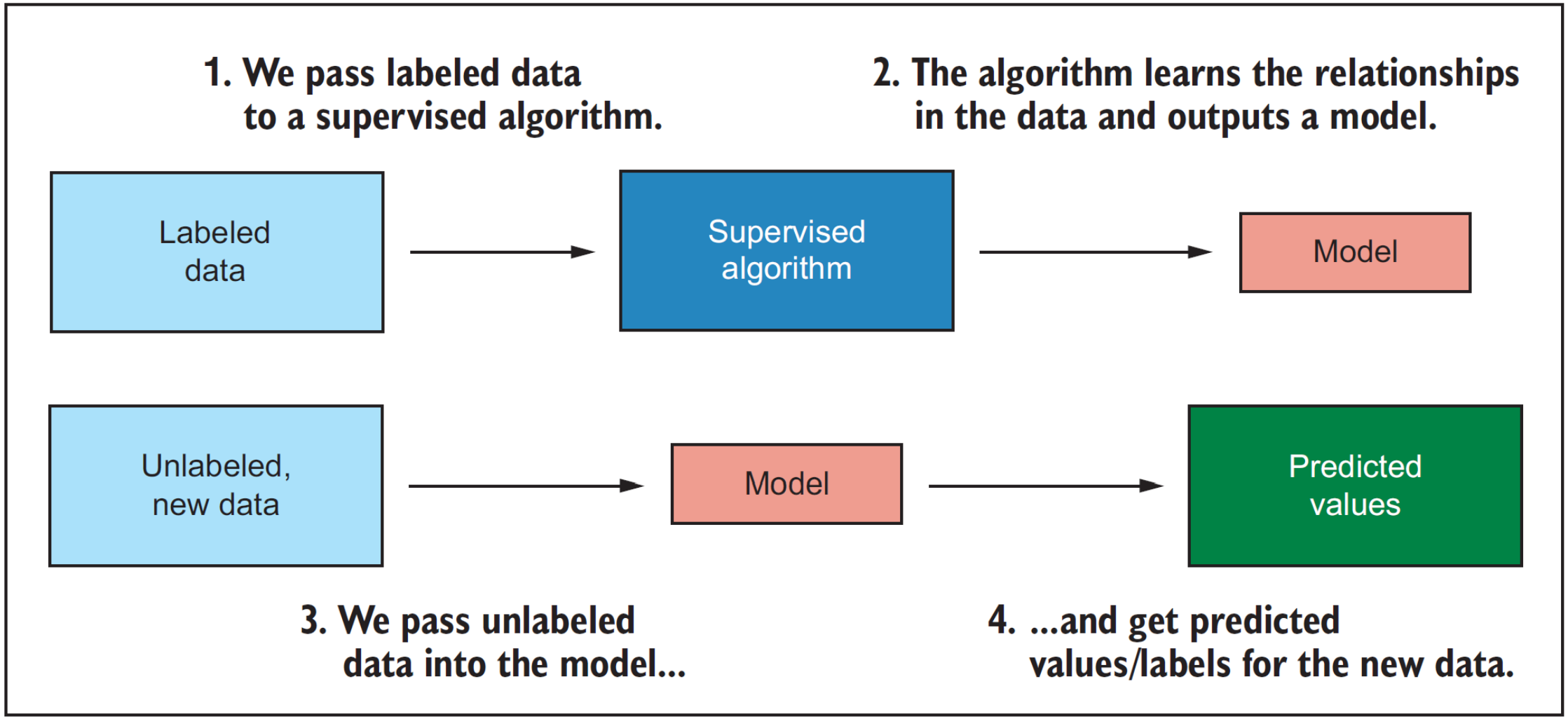
Unsupervised machine learning: Examples

University of Essex

- Clustering
- Dimension reduction
  - Principal Component Analysis
  - Scaling
    - Twitter followership network (Barbera, 2014)
- Text analytics
  - text scaling without pre-coded output
    - wordfish (Proksch and Slapin, 2008)
  - topic modeling
  - dictionary based sentiment classification (could be supervised)

Supervised machine learning, illustration

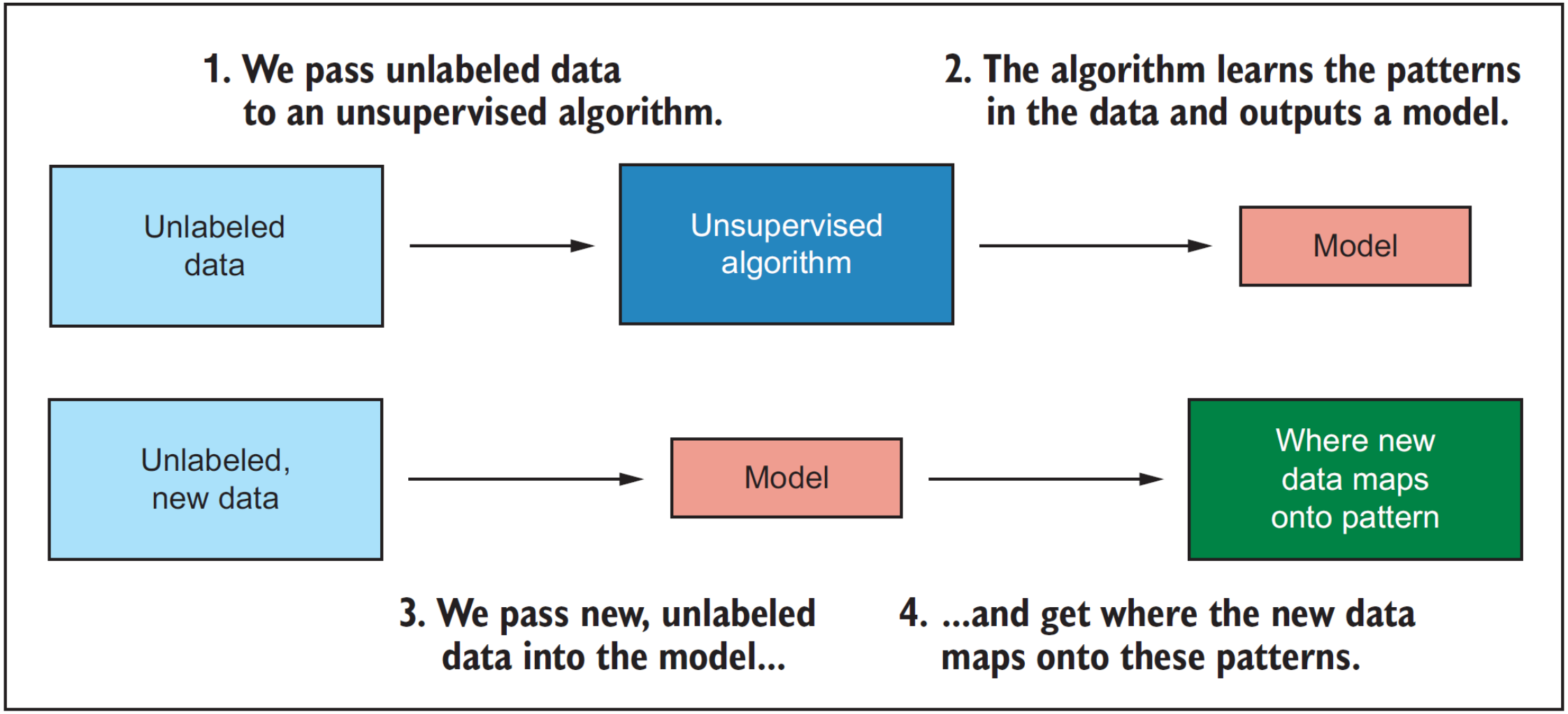
University of Essex



From Rhys, Machine Learning with R, the tidyverse, and mlr, 2020

Unsupervised machine learning, illustration

University of Essex



From Rhys, Machine Learning with R, the tidyverse, and mlr, 2020