UNIVERSITY OF WARSAW

WYDZIAŁ MATEMATYKI, INFORMATYKI I MECHANIKI

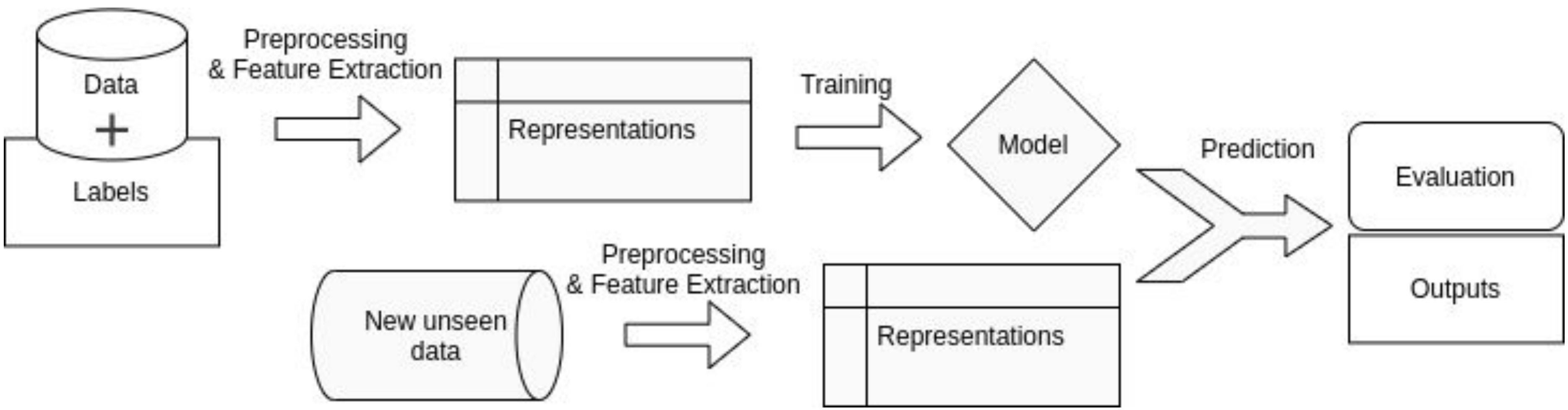
Interactive Machine Learning - Introduction

Andrzej Janusz
Daniel Kałuża

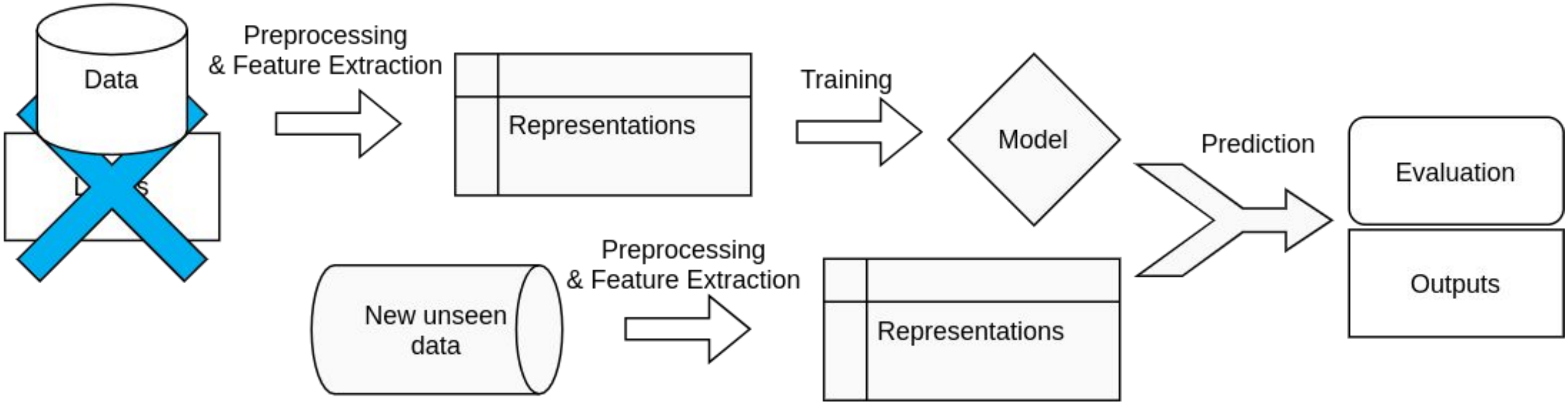
THE PLAN

- Introduction to Interactive Machine Learning
  - examples of ML process and related issues
  - various types of interaction in ML processes
  - real-life cases
- Selected aspects of Interactive ML
  - visual data mining, model monitoring tools, and typical data scientist work
  - active learning
- Topics discussed later in this course

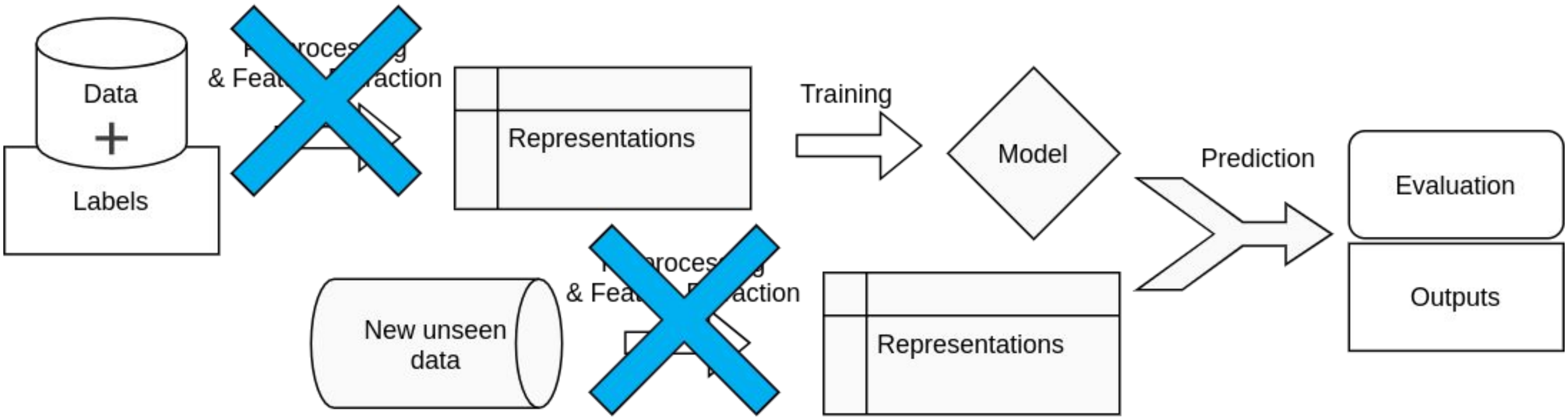
How does a typical ML process look like?



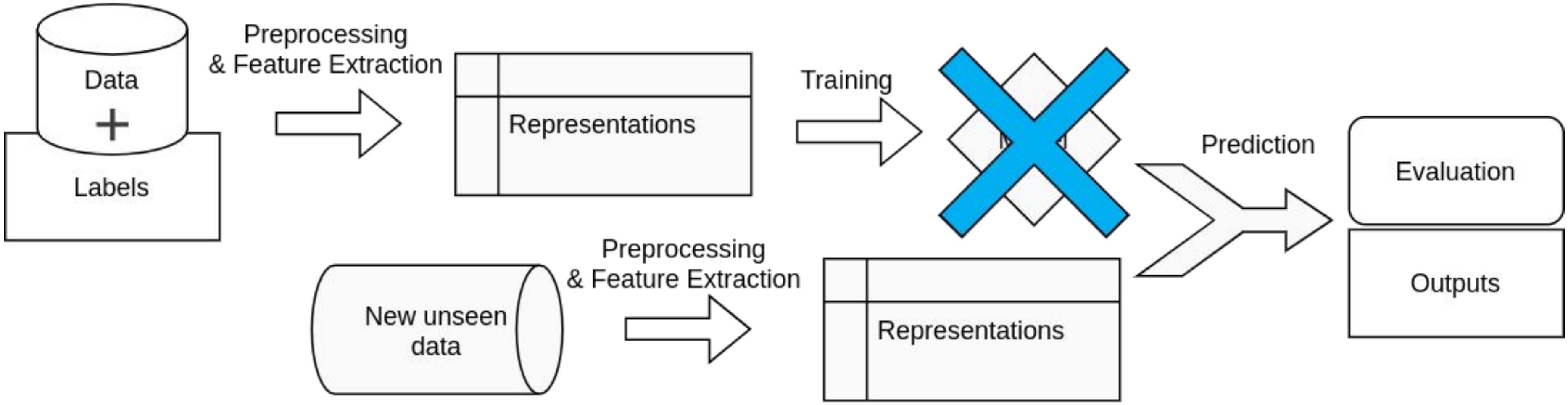
What if we don't have the labels?



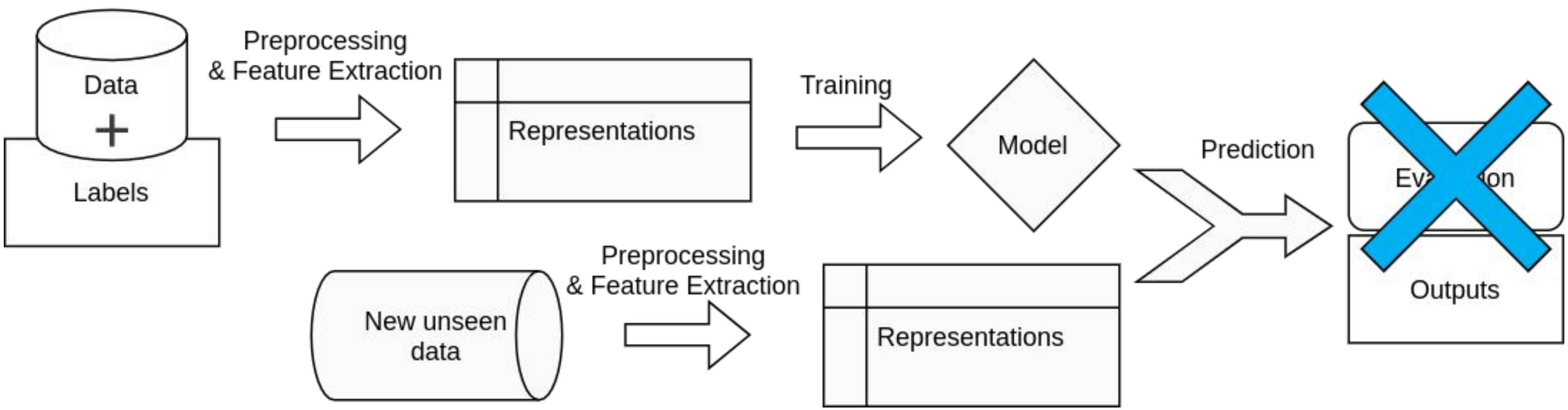
What if we don’t have a good data representation



What if we don't have sufficient resources for fully automatic model selection?



What if we don’t know what is the appropriate evaluation criteria?



Areas for the interaction in ML processes

- Acquisition of data labels
- Feature extraction + information fusion
- Model construction and monitoring
- Selection of evaluation criteria

Image: Freepik.com

Example 1 - life-sciences

- In life-sciences many experiments are costly and require a lot of processing time.
  - Many experiments have to be performed manually.
  - Experiments can often be parametrized.
- In many cases, ML models can be used to predict the expected experiment outcomes.
- Outcomes of an experiment may influence the parameter setup for the following experiments.

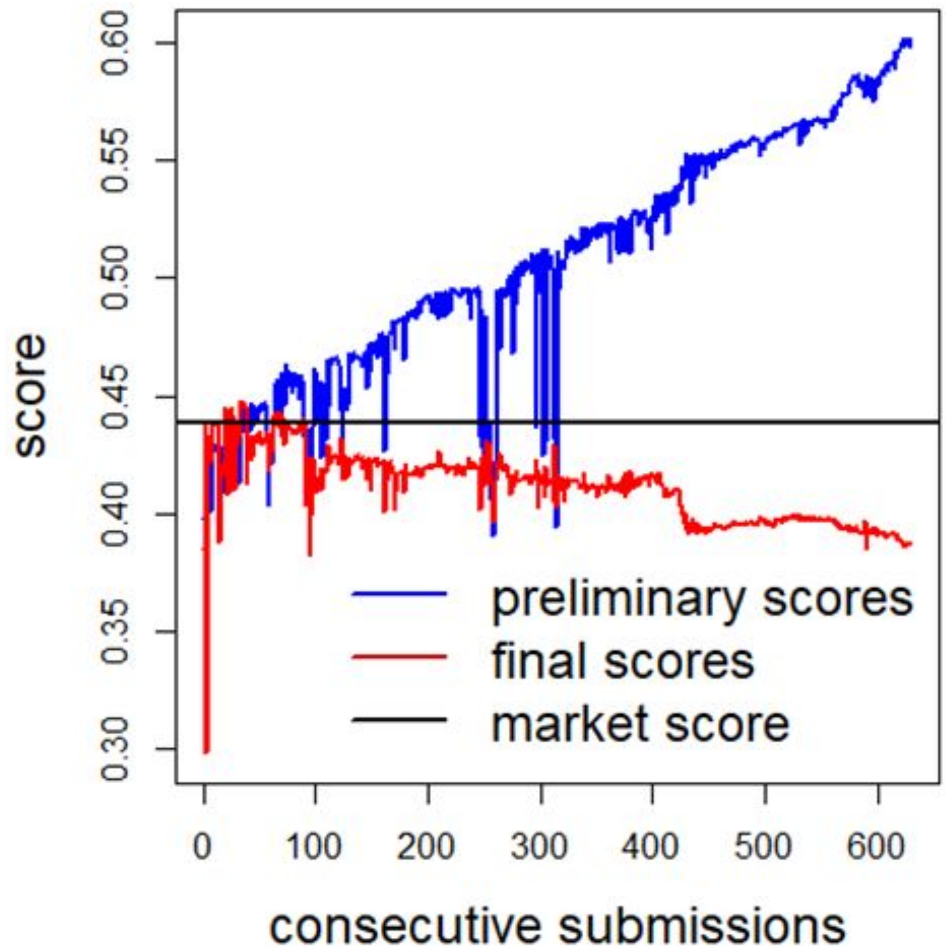
Image: Freepik.com

Example 2 - cybersecurity

- Companies that provide cybersecurity services have Security Operation Centers where human experts assess the severity of threat alerts.
- Multi-level systems:
  - automatically generated events,
  - expert rules for issuing alerts based on series of events,
  - automatic alert scoring systems,
  - human experts who assess the alert severity.
- Expert rules need to have very high sensitivity but generate many false alarms.
- Emerging attacks require adjustments to the expert rules.
- Human assessments can be used to improve specificity of the rules and quality of automatic alert scoring models.
- Scoring models may improve the efficiency of human experts!

Example 3 - a warning

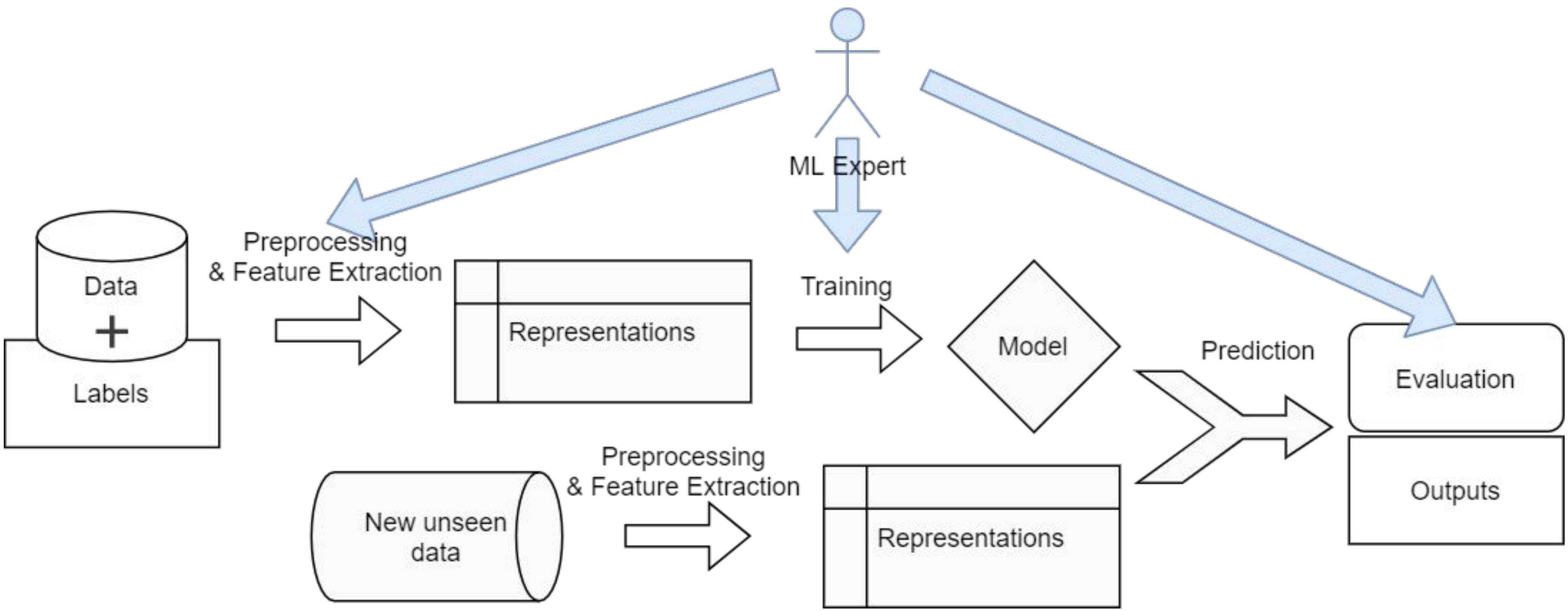
- The iterative model tuning can be harmful - the risk of overfitting to a validation set.
- The plot shows evaluation results of one team's submissions to a data mining competition organized at KnowledgePit (knowledgepit.ml).
- If the team stopped "improving" their solution after the 50th submission, they would have won.



Different types of interactions

- An analyst interacts with a data set - investigates and visualizes the data searching for correlations and potential issues.
- A data scientist designs data preprocessing steps - decides how to fill missing values, discretizes or encodes features, scales the data or performs the feature selection.
- A data scientist selects the most appropriate model for the task and performs hyper-parameter tuning.
- A deployed model is being monitored and its predictions are diagnosed.
- A model becomes outdated and needs adjustments.
- A group of experts labels the data to prepare a data set for the experiment or to construct an efficient prediction model.

Real-life ML processes



How to efficiently construct a good data set?

- Experts may label a random sample of data but...
  - what if the data is really big and...
  - the target distribution is skewed?
  - or the important concepts are rare?
- Can we do better than random sampling?
- Can we evaluate the quality of labels?

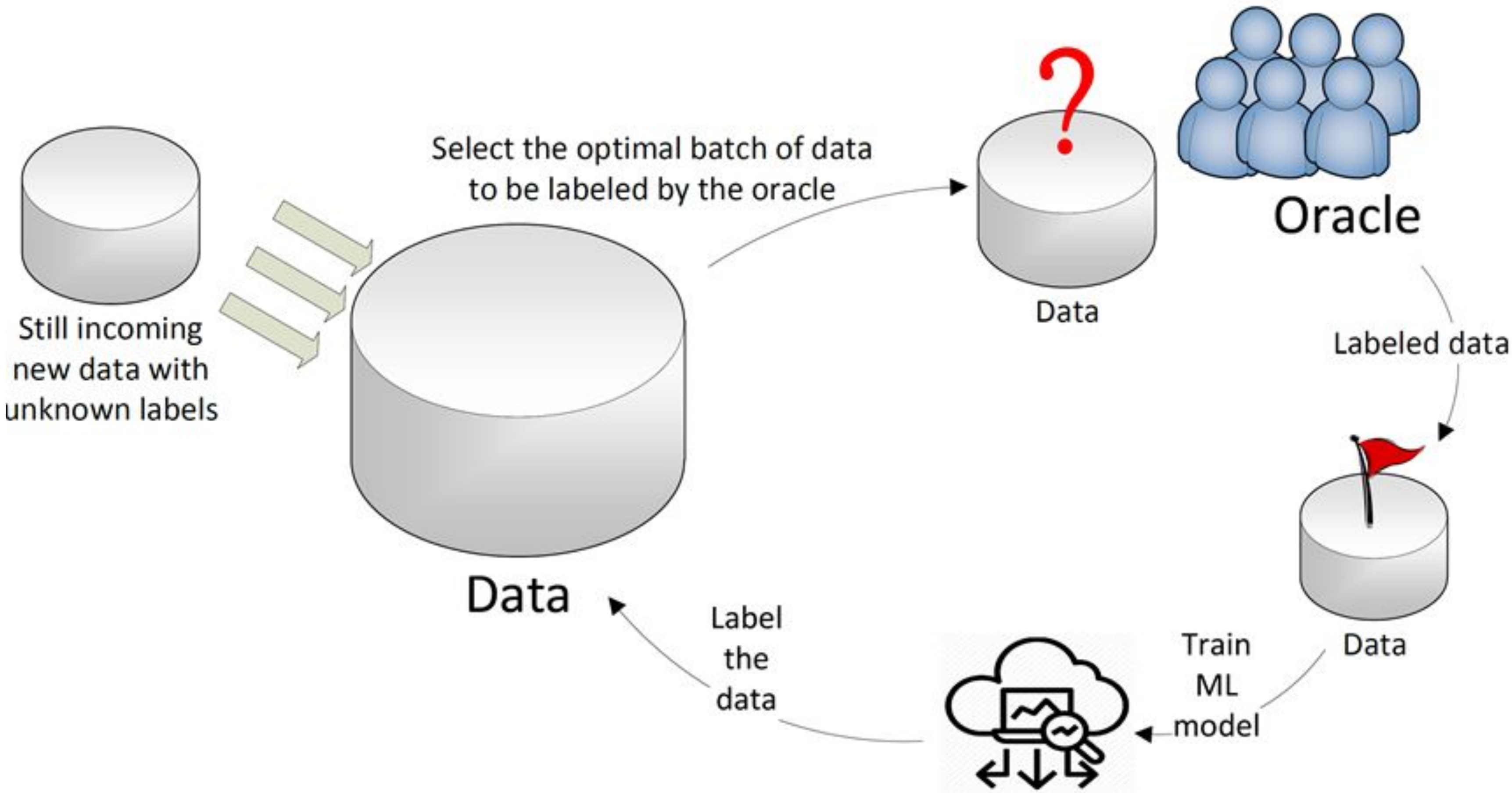
We will try to answer these questions in this course!

Active Learning

- Goal: obtain the best possible model with limited labeling capabilities, assuming that we can interact with the experts asking them to label indicated samples.
- How: iteratively query experts about labels for the most interesting/informative samples.
- Which samples are informative?
- Can we trust our experts?
- How do we update the model efficiently?

Image: Freepik.com

The active learning cycle



Adaptation to changing data

- What if the data distribution changes in time?
- Or completely new concepts/classes emerge?
- How can we detect such changes and adapt?

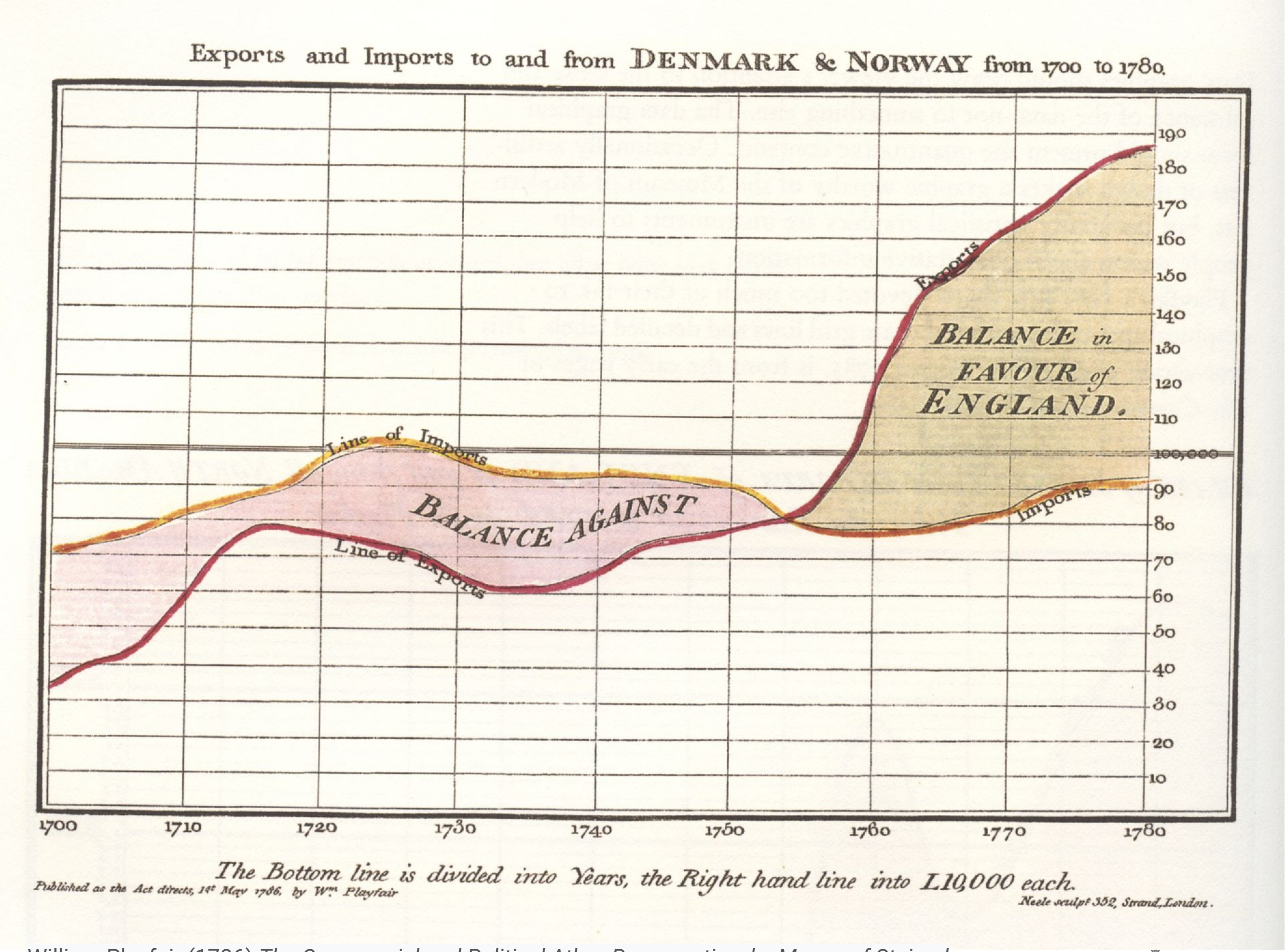
Image: Freepik.com

UNIVERSITY OF WARSAW

Visual data analytics

- Main focus areas:
  - Obtaining deep insights that support assessment, planning, and decision-making.
  - Representing data in ways that support visualization and analysis.
  - Supporting dissemination of the results of an analysis to clearly communicate information to a variety of audiences.
  - Enabling interaction techniques that allow users to explore and understand large amounts of information at once.
- Often, the selection of right data representation is the key.
- Theory of scientific data visualization: The Grammar of Graphics.
  - The most well-known implementation in the R package *ggplot2*.

Data visualization



William Playfair (1786) *The Commercial and Political Atlas: Representing, by Means of Stained Copper-Plate Charts, the Progress of the Commerce, Revenues, Expenditure and Debts of England during the Whole of the Eighteenth Century*. Source: Wikipedia (public domain)

Interactions can be (semi-)automatic

- Quality of input data and the performance of a prediction model should be monitored at all times.
  - XAI tools can be used to provide operators with insights
  - Alarms should be issued whenever the performance drops below a certain level.
- Pre-scheduled (automatic) vs. manual model updates.
- What can we say about prediction errors?
  - What can we learn from an advanced model audit?
  - Can such an audit be model-agnostic?

Summary of topics discussed in this course

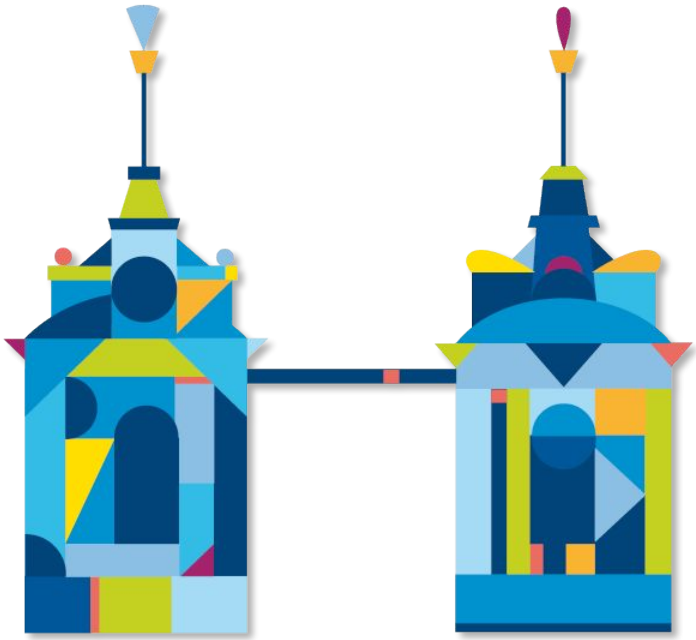
1. Active learning.
2. Selection of examples for effective model training.
   a. Uncertainty of models vs. example informativeness.
   b. Ensuring batch variety.
   c. Measures of the representativeness of examples.
   d. Optimizing the selection of examples for experts.
   e. Methods of achieving a consensus between voters on data labels.
3. Model updates and incremental learning.
4. Drift of concepts and interactive adaptation of models.
5. Unsupervised learning, semi-supervised learning, and self-supervised learning.
6. Interactive discovery of anomalies in data.
7. Counterfactual explanations - the prescriptive analysis.
8. Other topics (depending on available time):
   a. Interactive feature engineering.
   b. Visual data mining.
   c. Selected issues in the field of life-long learning.

The rules

- Our course in Moodle: https://moodle.mimuw.edu.pl/course/view.php?id=1403
  Passcode: **2W+)rt_&**
- Two lab projects in a form of data challenges:
  - Second half of March.
  - Second half of April.
- Presentations of selected scientific papers.
- Oral exam in September.

Literature:

1. B. Settles. Active Learning Literature Survey. Computer Sciences Technical Report 1648, University of Wisconsin–Madison, (2010).
2. R.D. King, K.E. Whelan, F.M. Jones, P.G. Reiser, C.H. Bryant, S.H. Muggleton, D.B. Kell, and S.G. Oliver. Functional genomic hypothesis generation and experimentation by a robot scientist. Nature, 427(6971):247–52, 2004.
3. R.D. King, J. Rowland, S.G. Oliver, M. Young, W. Aubrey, E. Byrne, M. Liakata, M. Markham, P. Pir, L.N. Soldatova, A. Sparkes, K.E. Whelan, and A. Clare. The automation of science. Science, 324(5923):85–89, 2009.
4. L. Wilkinson. 2005. The Grammar of Graphics (Statistics and Computing). Springer-Verlag, Berlin, Heidelberg.
5. J. Bosser, E. Sorstadius and M. Chehreghani, "Model-Centric and Data-Centric Aspects of Active Learning for Deep Neural Networks," in 2021 IEEE International Conference on Big Data (Big Data), Orlando, FL, USA, 2021 pp. 5053-5062.
6. T. Mitchell et al. 2018. Never-ending learning. Commun. ACM 61, 5 (May 2018), 103–115.



QUESTIONS OR COMMENTS?

a.janusz@mimuw.edu.pl

or

d.kaluza@mimuw.edu.pl