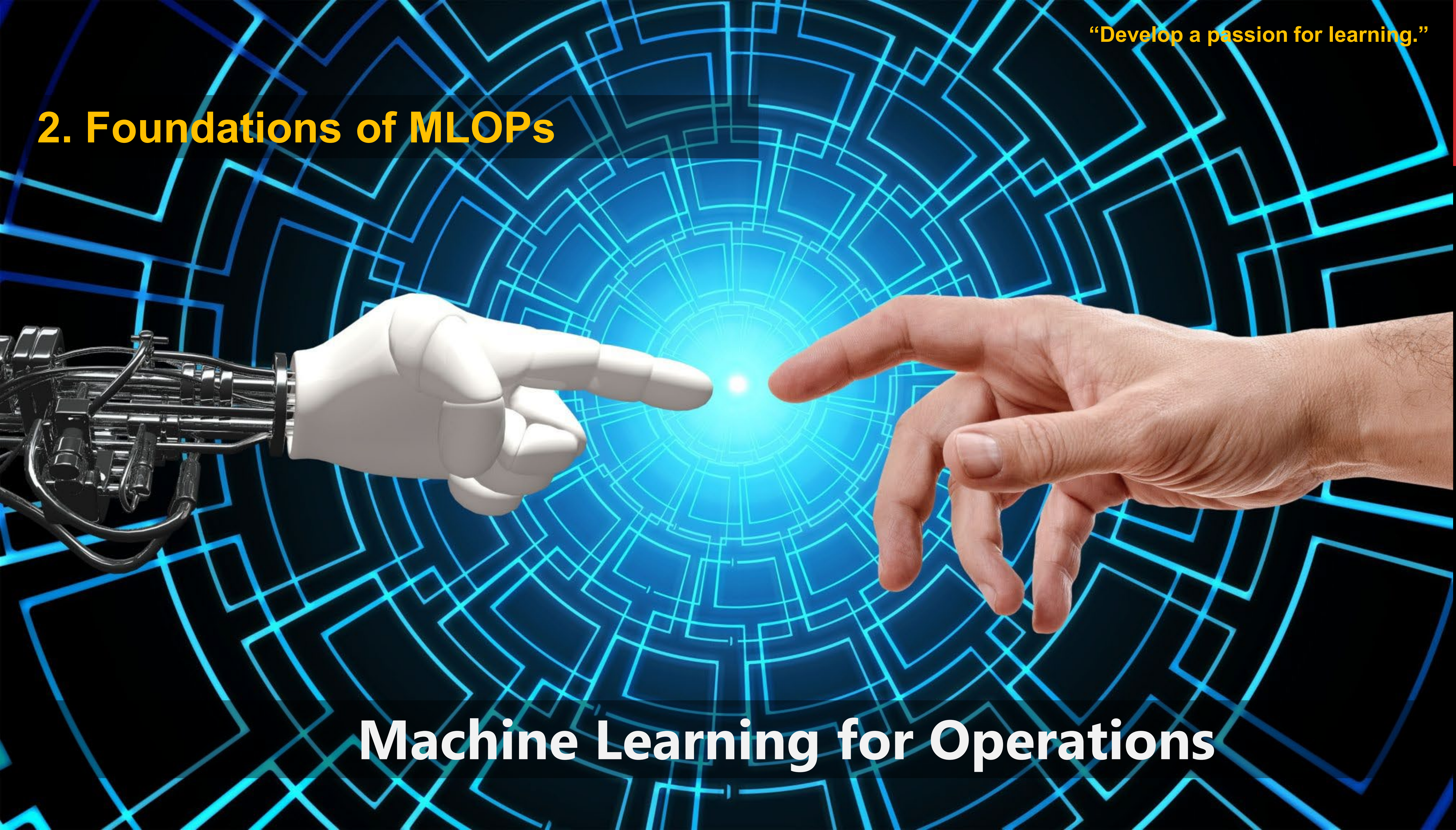
“Develop a passion for learning.”

2. Foundations of MLOPs

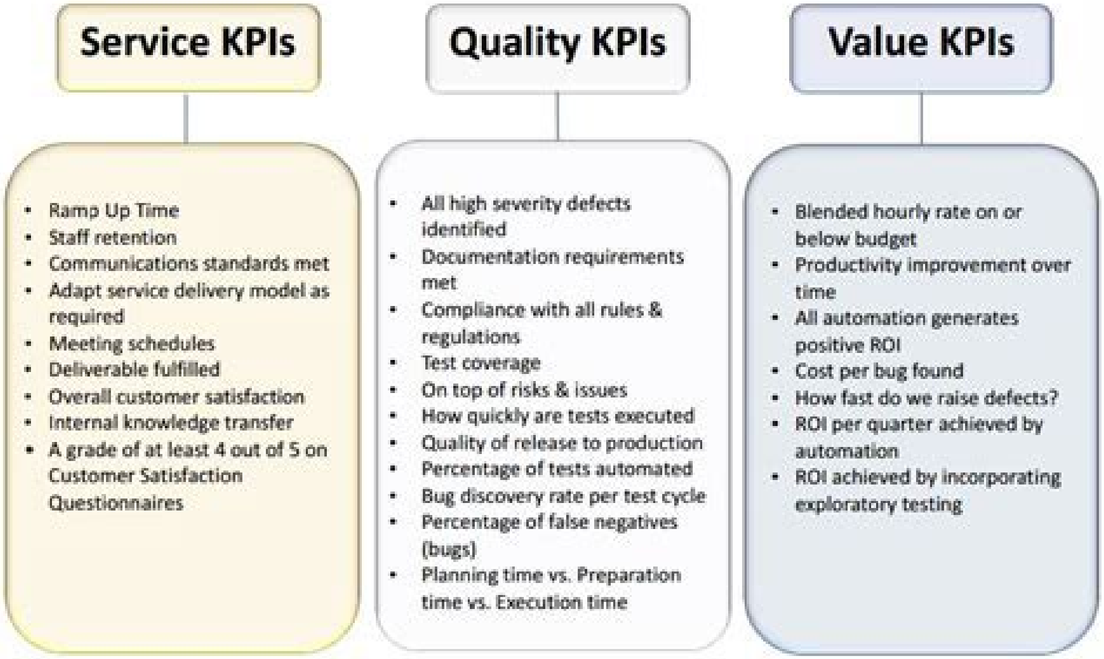


Machine Learning for Operations

A Primer on Machine Learning

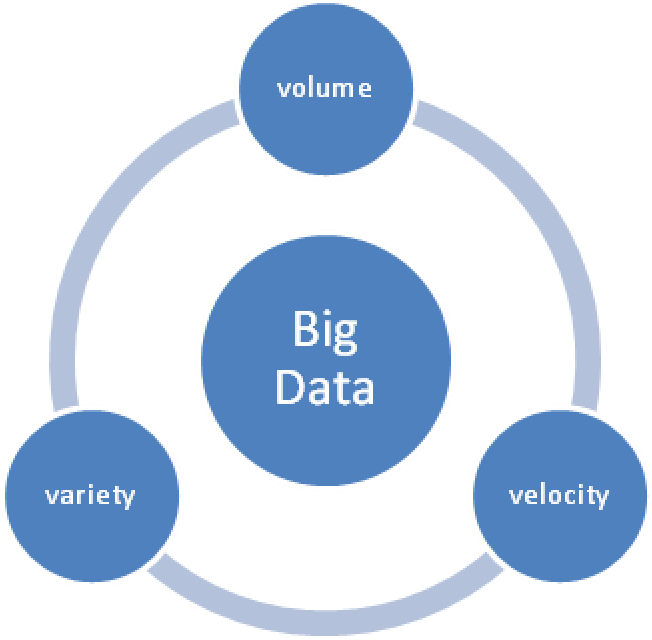
Understanding the Business

- What does the business need?
  - What are the problems that need to be solved?
  - Is there enough information to state the problem?
  - What are the specific outcomes that are needed?
- Service Level Agreements
  - What are the quality levels needed?
  - How do we measure outcomes?
  - What is the necessary business outcome?
  - How will it be measured?
- Key Performance Indicators
  - What needs to be measured and why?
  - How do the KPIs relate to SLAs?



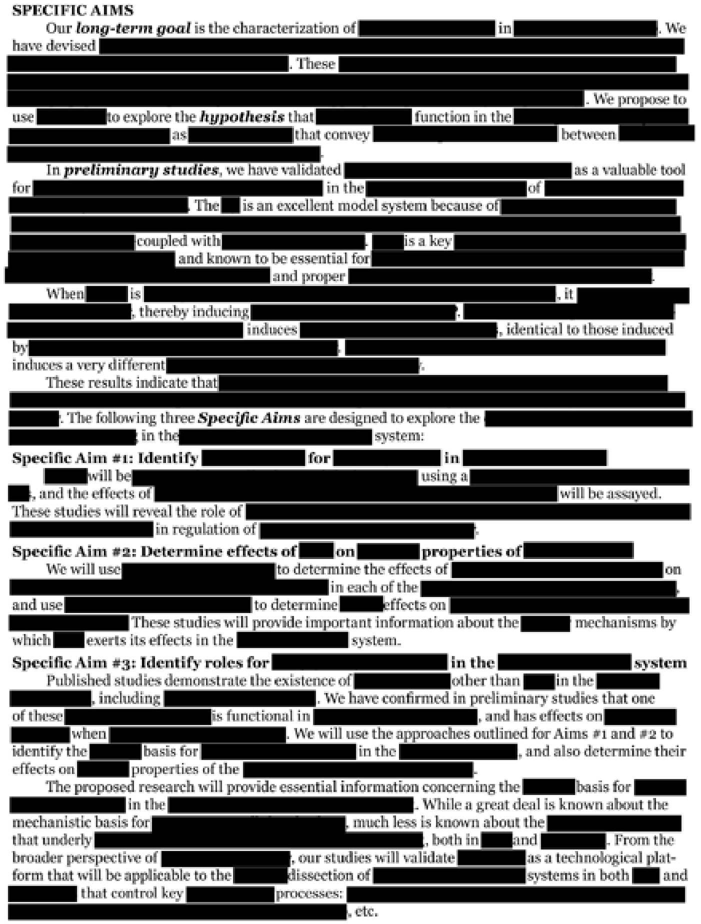
Data Sources and Exploratory Data Analysis

- Key questions to be answered:
  - What relevant datasets are available?
  - Is this data sufficiently accurate and reliable?
  - How can stakeholders get access to this data?
  - What data properties (known as features) can be made available by combining multiple sources of data?
  - Will this data be available in real time?
  - Is there a need to label some of the data with the “ground truth” that is to be predicted, or does unsupervised learning make sense? If so, how much will this cost in terms of time and resources?
  - What platform should be used?
  - How will data be updated once the model is deployed?
  - Will the use of the model itself reduce the representativeness of the data?
  - How will the KPIs, which were established along with the business objectives, be measured?



Data Constraints

- Key questions to be answered:
  - Can the selected datasets be used for this purpose?
  - What are the terms of use?
  - Is there personally identifiable information (PII) that must be redacted or anonymized?
  - Are there features, such as gender, that legally cannot be used in this business context?
  - Are minority populations sufficiently well represented that the model has equivalent performances on each group?



SPECIFIC AIMS

Our ***long-term goal*** is the characterization of [redacted] in [redacted]. We have devised [redacted]. These [redacted]. We propose to use [redacted] to explore the ***hypothesis*** that [redacted] function in the [redacted] as [redacted] that convey [redacted] between [redacted].

In ***preliminary studies***, we have validated [redacted] as a valuable tool for [redacted] in the [redacted] of [redacted]. The [redacted] is an excellent model system because of [redacted] coupled with [redacted]. [redacted] is a key [redacted] and known to be essential for [redacted] and proper [redacted].

When [redacted] is [redacted], it [redacted], thereby inducing [redacted]. [redacted] induces [redacted], identical to those induced by [redacted]. [redacted] induces a very different [redacted].

These results indicate that [redacted]. The following three ***Specific Aims*** are designed to explore the [redacted] in the [redacted] system:

Specific Aim #1: Identify [redacted] for [redacted] in [redacted]

[redacted] will be [redacted] using a [redacted], and the effects of [redacted] will be assayed. These studies will reveal the role of [redacted] in regulation of [redacted].

Specific Aim #2: Determine effects of [redacted] on [redacted] properties of [redacted]

We will use [redacted] to determine the effects of [redacted] on [redacted] in each of the [redacted]. and use [redacted] to determine [redacted] effects on [redacted]. These studies will provide important information about the [redacted] mechanisms by which [redacted] exerts its effects in the [redacted] system.

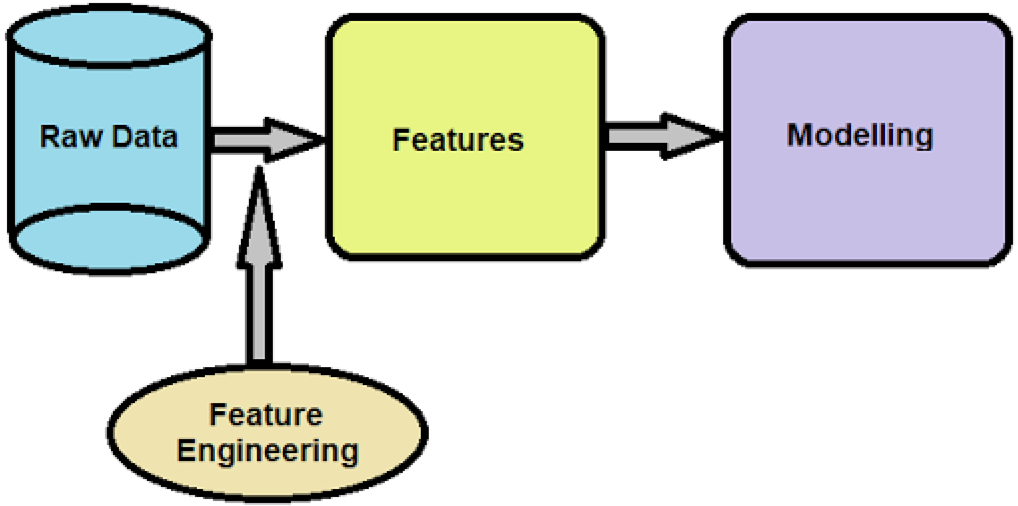
Specific Aim #3: Identify roles for [redacted] in the [redacted] system

Published studies demonstrate the existence of [redacted] other than [redacted] in the [redacted], including [redacted]. We have confirmed in preliminary studies that one of these [redacted] is functional in [redacted], and has effects on [redacted] when [redacted]. We will use the approaches outlined for Aims #1 and #2 to identify the [redacted] basis for [redacted] in the [redacted], and also determine their effects on [redacted] properties of the [redacted].

The proposed research will provide essential information concerning the [redacted] basis for [redacted] in the [redacted]. While a great deal is known about the mechanistic basis for [redacted], much less is known about the [redacted] that underly [redacted], both in [redacted] and [redacted]. From the broader perspective of [redacted], our studies will validate [redacted] as a technological platform that will be applicable to the [redacted] dissection of [redacted] systems in both [redacted] and [redacted] that control key [redacted] processes: [redacted], etc.

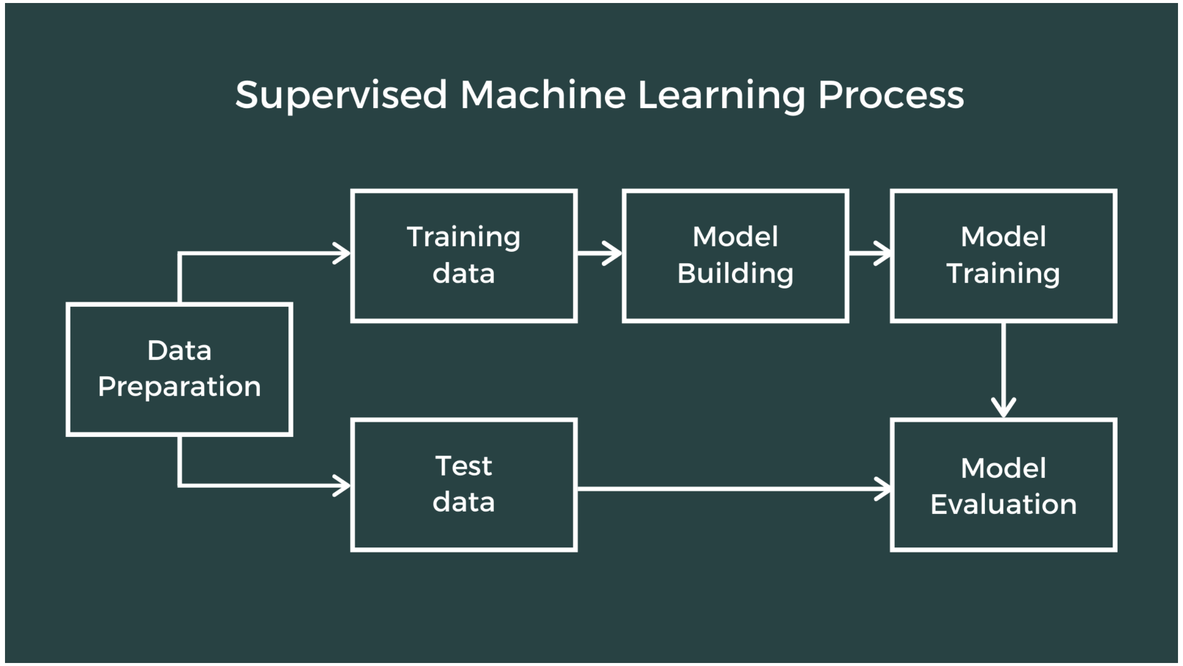
Feature Engineering and Selection

- Critical part of model development
- Poor feature selection impacts MLOps in a variety of ways
  - The model can become more and more expensive to compute.
  - More features require more inputs and more maintenance down the line.
  - More features mean a loss of some stability.
  - The sheer number of features can raise privacy concerns.



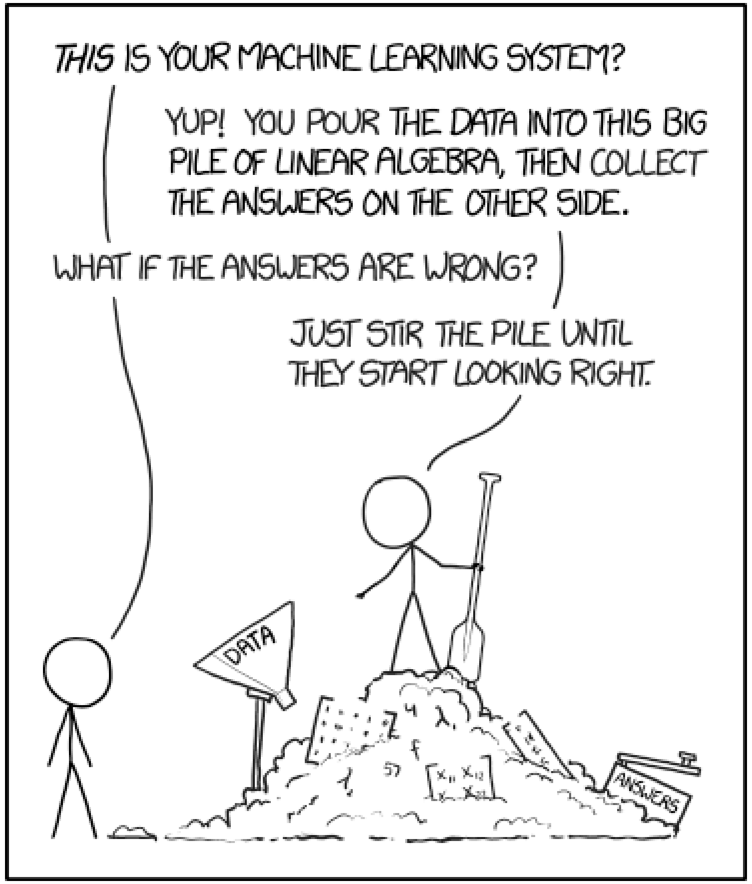
Training and Evaluation

- Standard methodology in ML
- For MLOps
  - Is the training data realistic?
- We train a model on fixed data
  - Tune results by algorithm choice
  - Tune results by feature selection
  - Tune results with hyperparameters
- Overfitting is real concern
- Setting evaluation criteria is important
  - Moderate success on wide ranges of training data versus high success on narrow sets of training data



Reproducibility

- Supported by good CI practices
- What was created in dev can be reproduced in prod
- Ability to roll back to previous versions
- Requires version control of ALL assets used in developing the model



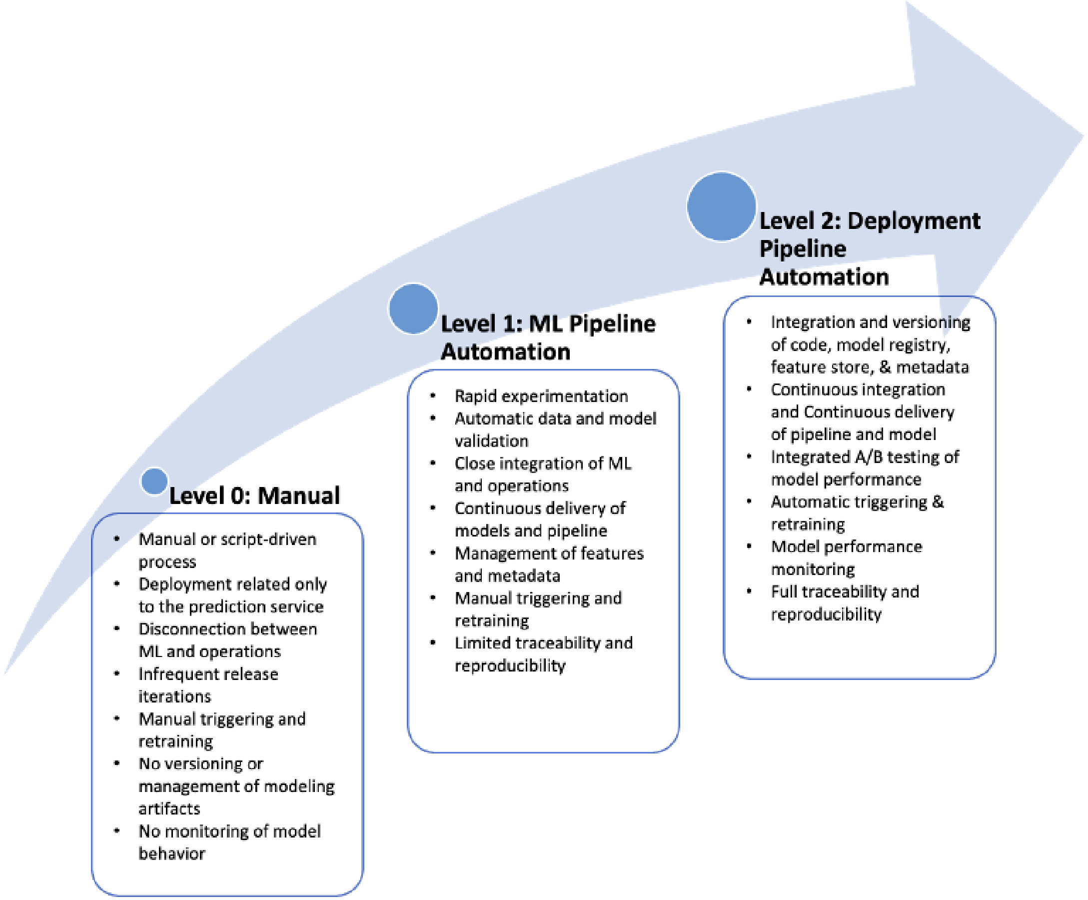
Responsible AI

- Results need to be explainable
- Mitigate uncertainty and help prevent unintended consequences
- Uses:
  - Partial dependence plots, which look at the marginal impact of features on the predicted outcome
  - Subpopulation analyses, which look at how the model treats specific subpopulations and that are the basis of many fairness analyses
  - Individual model predictions, such as Shapley values, which explain how the value of each feature contributes to a specific prediction
  - What-if analysis, which helps the ML model user to understand the sensitivity of the prediction to its inputs

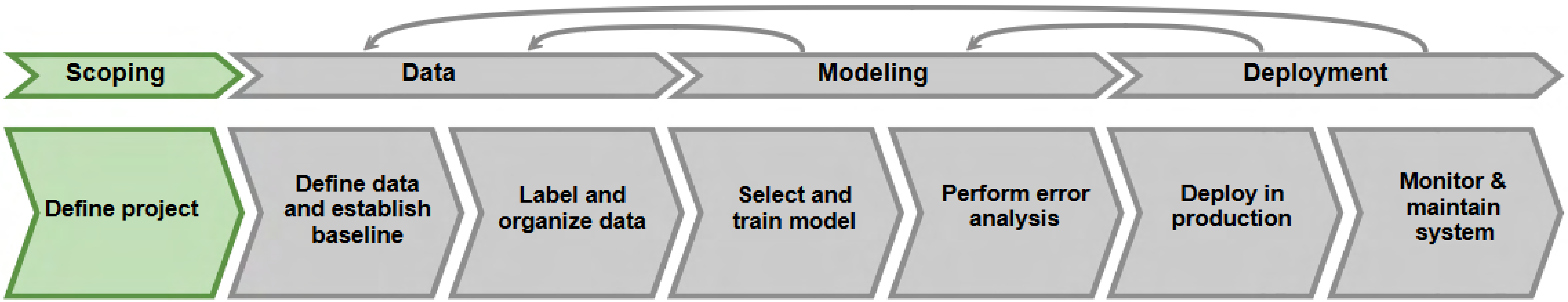
MLOps Maturity Levels

Maturity Levels

- Various attempts to quantify the organization's level of expertise in using MLOps
  - https://www.eqengineered.com/insights/mlops-maturity-levels
  - https://docs.microsoft.com/en-us/azure/architecture/example-scenario/mlops/mlops-maturity-model

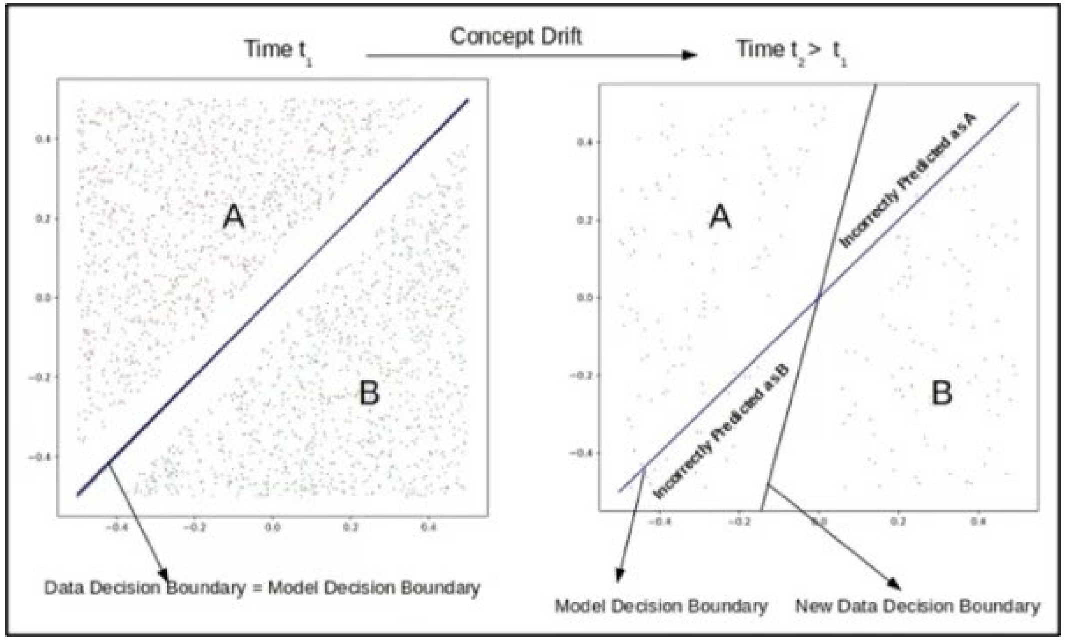


Model Lifecycle



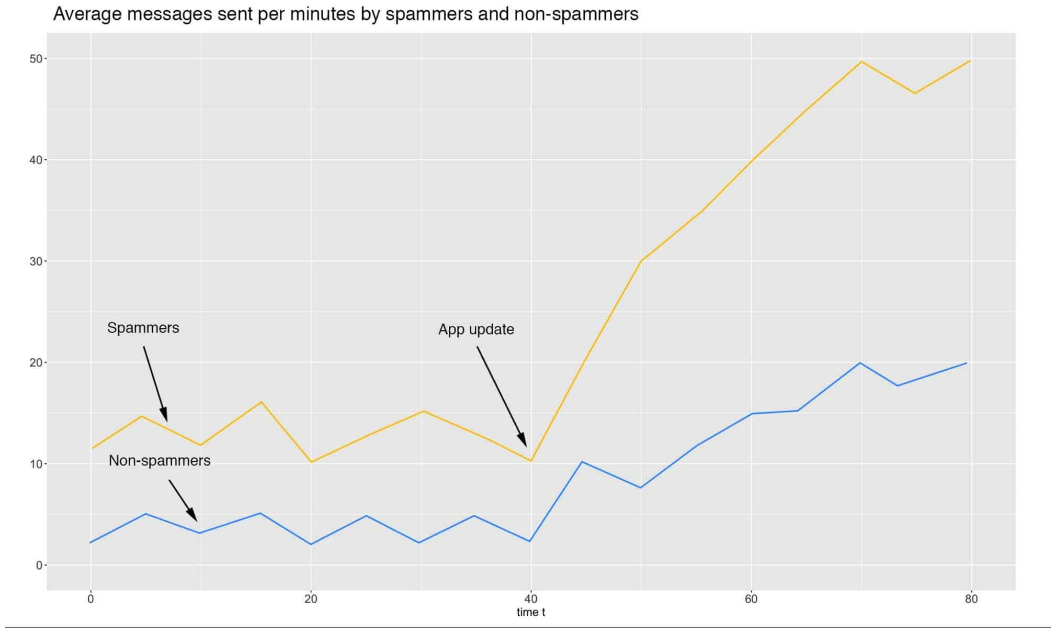
Concept Drift

- ML models are trained assuming a certain conceptual relationship between inputs and predictions
  - When that relationship changes in the real world, we have concept drift
  - The model no longer describes the actual relationships
  - For example, floor area is no longer a factor in home prices.
  - Indicates a change in the decision boundary
- Concept drift can be managed only by retraining the model



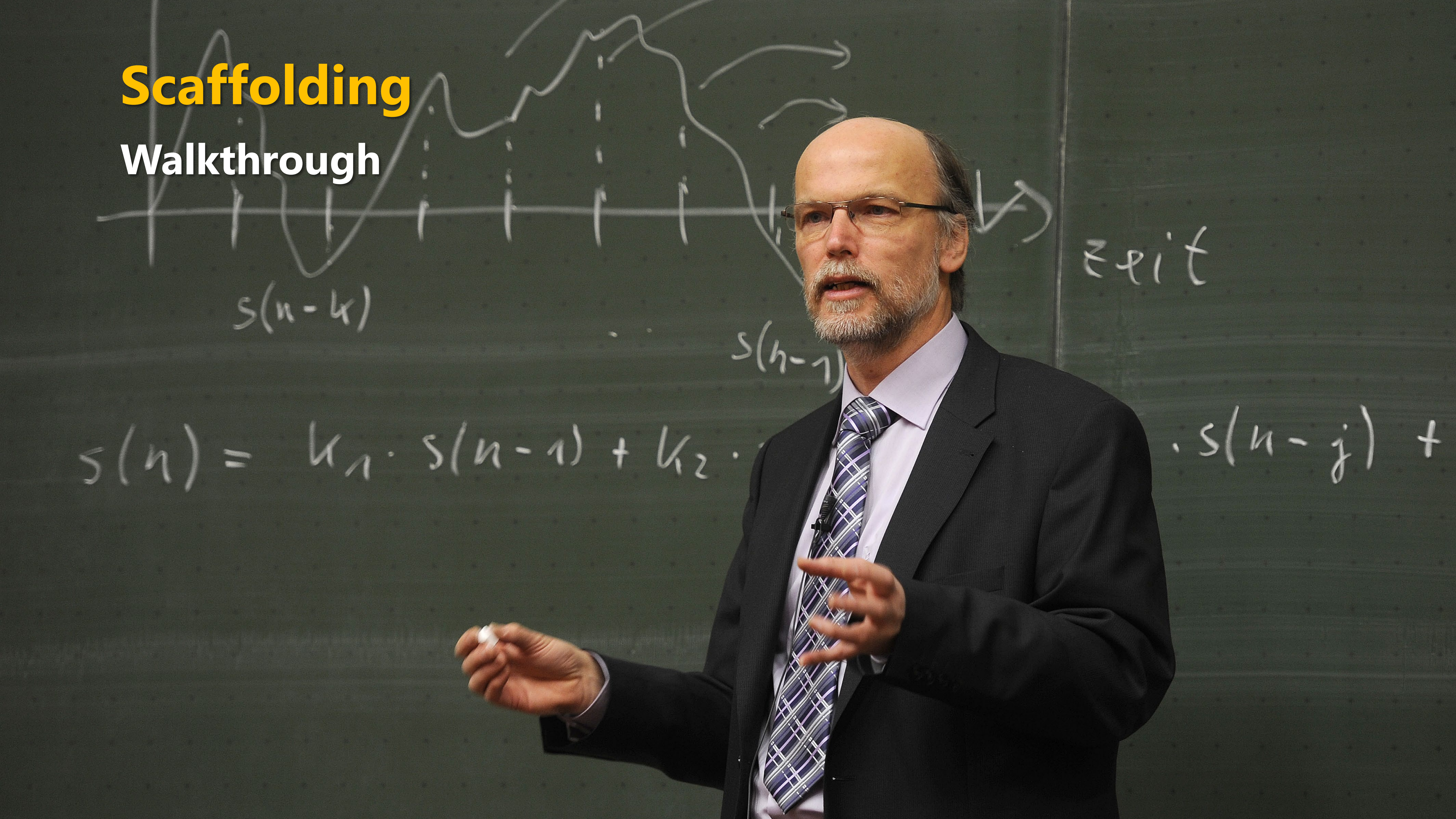
Data Drift

- Key ML assumption
  - Future is predictable from the past
- Data drift is when
  - The distribution of the input data has changed
- Temporal shifts
  - The population changes over time – changing demographics for example
  - Features in the model no longer correspond to features in the data
- Spatial shifts
  - The model is deployed in a different population than it was trained for
  - US models deployed in China



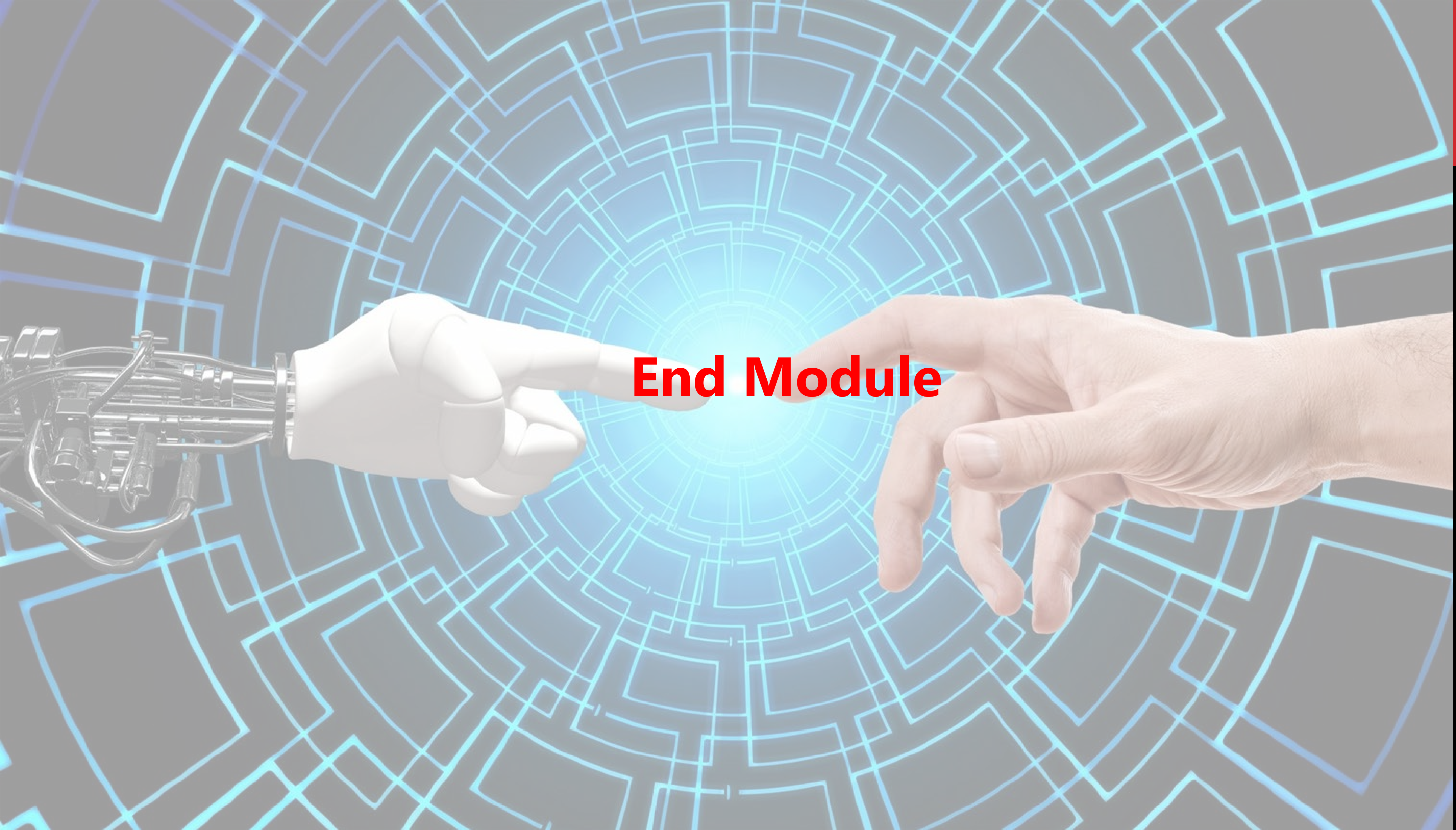
Scaffolding

Walkthrough



Scaffolding

Lab 3


End Module