TOWARDS FEATURE-BASED ANALYSIS OF THE MACHINE LEARNING DEVELOPMENT LIFECYCLE

ESEC/FSE IVR 2023

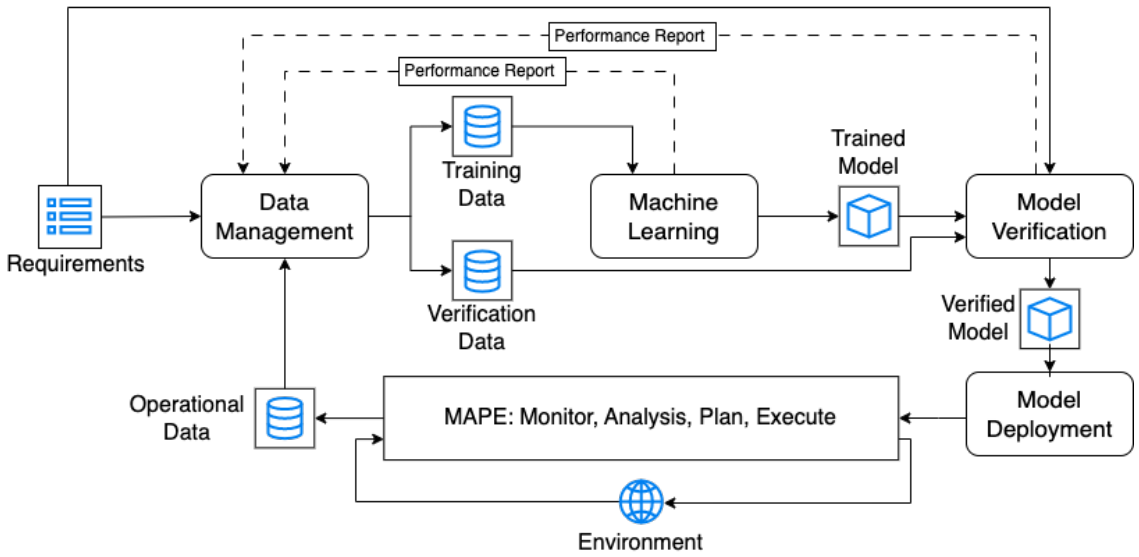
UNIVERSITY OF TORONTO

Boyue Caroline Hu

Marsha Chechik

MOTIVATION

ML development lifecycle
[Ashmore et. al, 2021]



Application of ML models in larger systems requires an in-depth analysis of each **ML development lifecycle** stage:

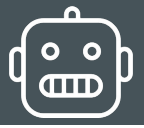
- Compliance between input and output artifacts in each stage
- Consistency of artifact usage across stages

Challenges:

- Missing detailed specifications
- Each stage requires specific expertise

Rob Ashmore, Radu Calinescu, and Colin Paterson. 2021. Assuring the Machine Learning Lifecycle: Desiderata, Methods, and Challenges. ACM Comput. Surv. 54, 5, Article 111 (2021), https://doi.org/10.1145/3453444

FEATURES



Train or airplane?

"Given a distribution of data, ML models aim to extract and learn generalizable features for the task they are performing." [Hinton, 2014]

Def.: features are high-level abstractions of desired functionalities, model behaviour, and data

Features in ML development lifecycle stages:

- *Model Verification*: check intended behaviour
  - e.g., background and performance
- *Data Management*: data augmentation
  - e.g., balance images with different background during training
- …

Example: the background feature

	train	airplane
sky		
concrete		

Geoffrey Hinton. 2014. Where do features come from? Cognitive science 38, 6 (2014), 1078–1101. https://doi.org/10.1111/cogs.12049

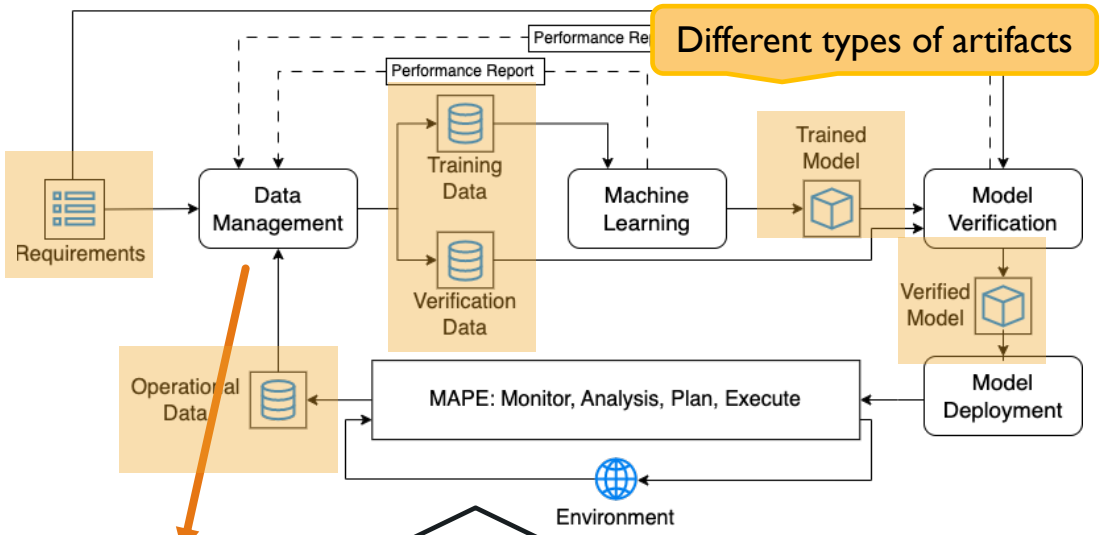
OUR VISION

Features present the desired abstraction to

- Support communication between the ML development lifecycle stages
- Support communication between stakeholders with different training background
- Enable incremental development and validation

Our vision (feature-based analysis of the ML Development Lifecycle): use features across all stages of the lifecycle to support software engineers to ...

1) … enable input-output compliance per stage
2) … analyze interactions between different stages
3) … enable modular development of artifacts



Example compliance for *Data Management*
Input: requirements
Output: data (training and verification)
Feature analysis: data with different background are fairly represented
Modular analysis: reuse the background feature for bird VS cars

FEATURE-BASED ANALYSIS OF THE ML DEVELOPMENT LIFECYCLE

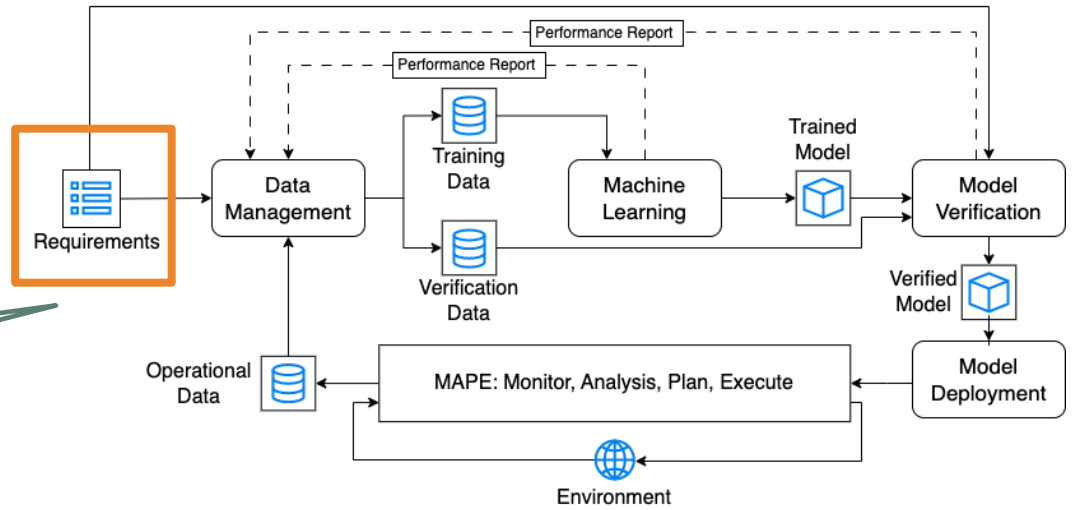
Goal: ensure compliance of features between the input and output artifacts

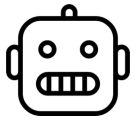
- i.e., is each feature of the input artifact sufficiently represented in the output artifact?

Start with system-level description to obtain requirements w.r.t. features and their interactions

Example:

- Specify expected performance of the ML w.r.t different backgrounds
- Specify data composition of different backgrounds



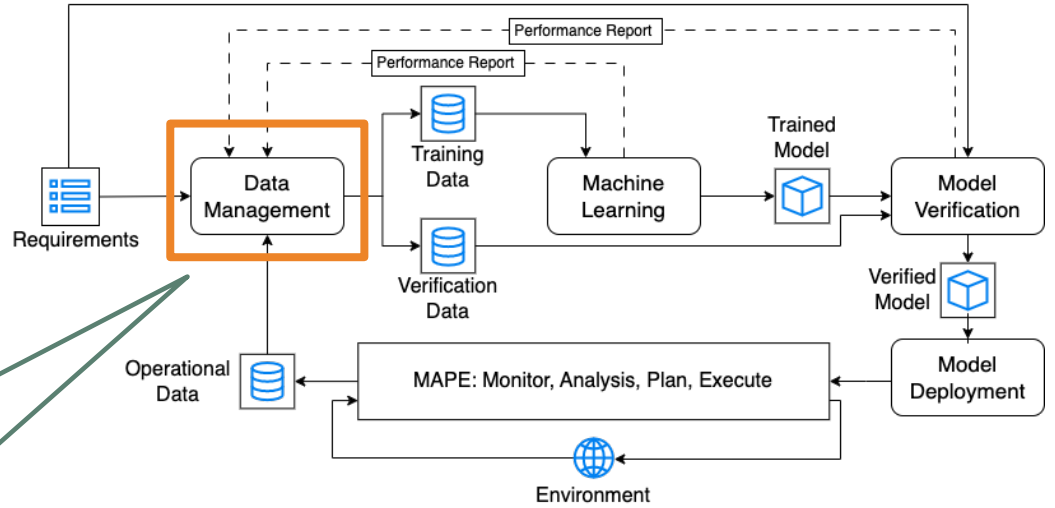


Train or airplane?

FEATURE-BASED ANALYSIS OF THE ML DEVELOPMENT LIFECYCLE

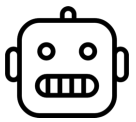
Goal: ensure compliance of features between the input and output artifacts

- i.e., is each feature of the input artifact sufficiently represented in the output artifact?



Data Management stage:

Goal: check data sufficiency for a specified feature
Example: Check if data has a fair distribution of both the sky background and the concrete background

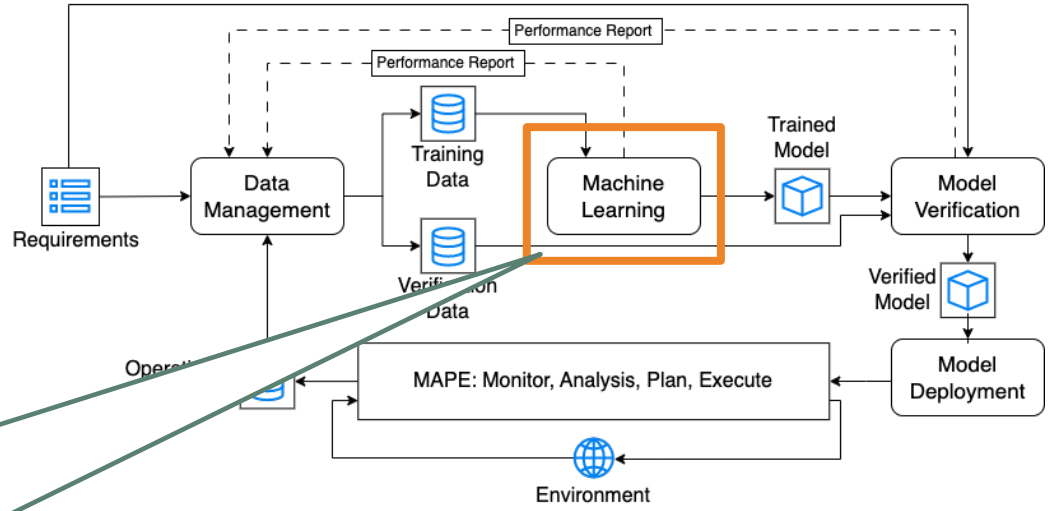


Train or airplane?

FEATURE-BASED ANALYSIS OF THE ML DEVELOPMENT LIFECYCLE

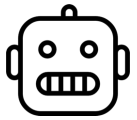
Goal: ensure compliance of features between the input and output artifacts

- i.e., is each feature of the input artifact sufficiently represented in the output artifact?



Machine Learning stage:

Goal: check that model achieves satisfactory performance for features represented in the data.
Example: measure model accuracy for different backgrounds.



Train or airplane?

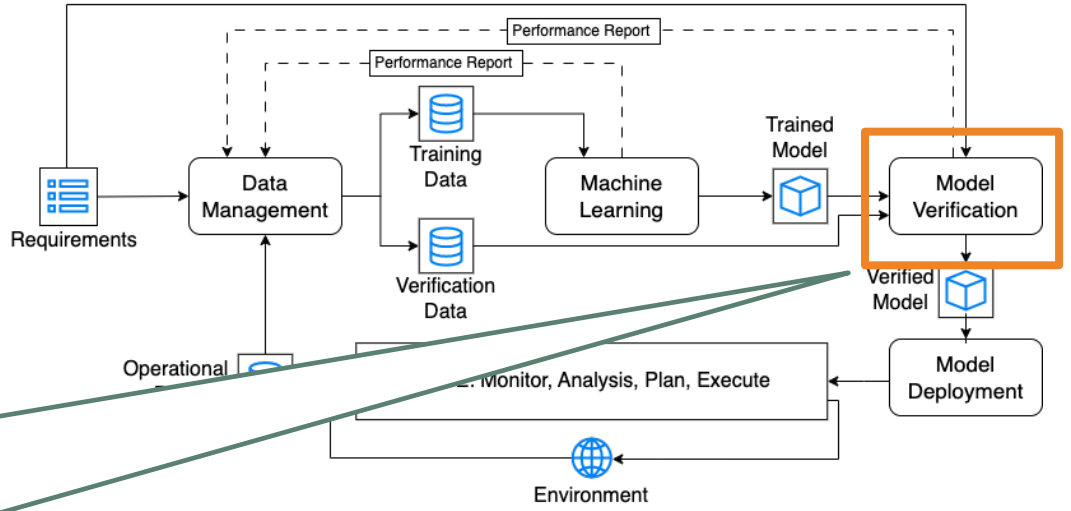
FEATURE-BASED ANALYSIS OF THE ML DEVELOPMENT LIFECYCLE

Goal: ensure compliance of features between the input and output artifacts

- i.e., is each feature of the input artifact sufficiently represented in the output artifact?

Model Verification stage:

Goal: for each feature specified in the requirements, check whether ML reached desired performance
Example: perform rigorous testing of model performance with different backgrounds

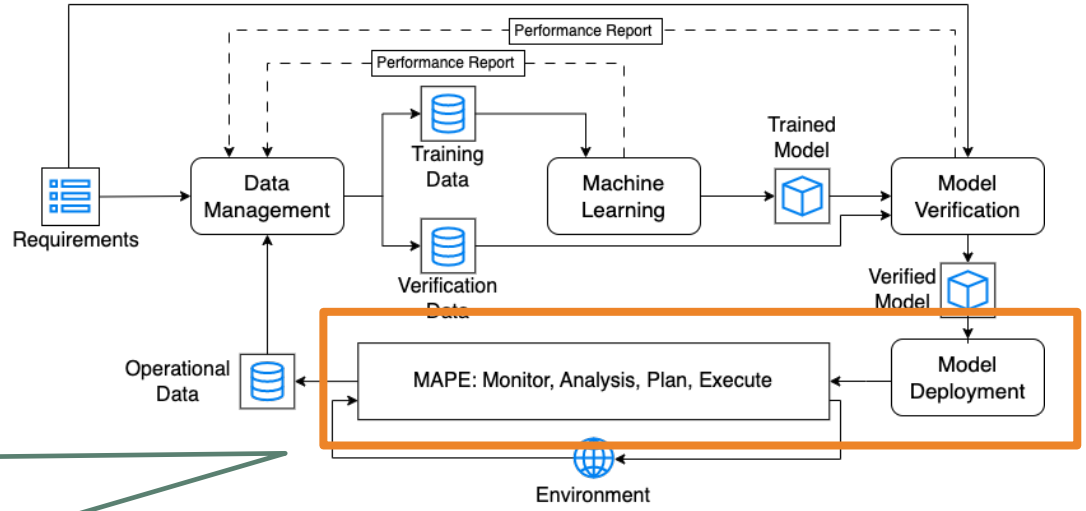


Train or airplane?

FEATURE-BASED ANALYSIS OF THE ML DEVELOPMENT LIFECYCLE

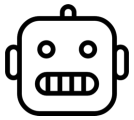
Goal: ensure compliance of features between the input and output artifacts

- i.e., is each feature of the input artifact sufficiently represented in the output artifact?



Model Deployment and *MAPE:*

Goal: monitor ML model behaviour per feature during deployment
Example: monitor accuracy for different backgrounds during deployment

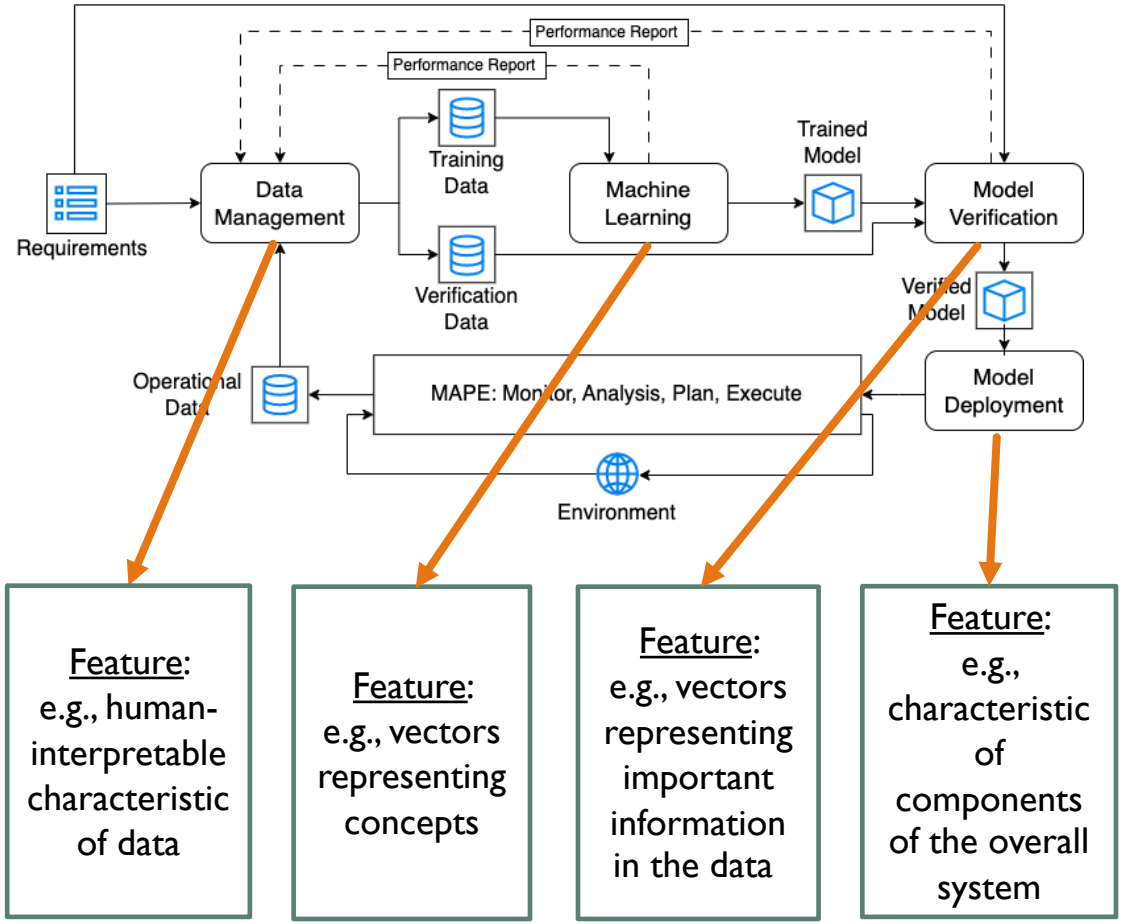


Train or airplane?

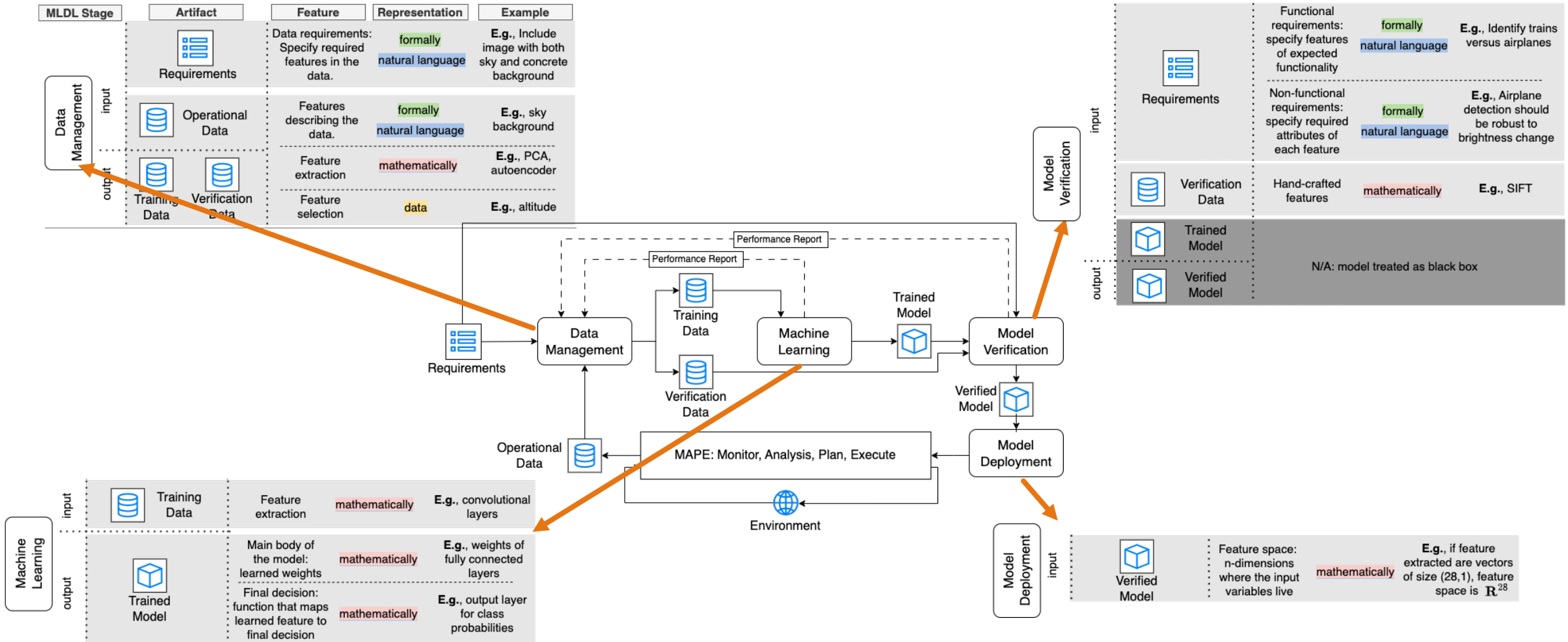
TOWARDS OUR VISION

Different stages of ML development lifecycle involve different expertise, definitions, representations and usages of features

<u>Built a taxonomy of features</u>
by systematically
identifying definitions of features
used by state-of-the-art methods in
each lifecycle stage



TAXONOMY OF FEATURES



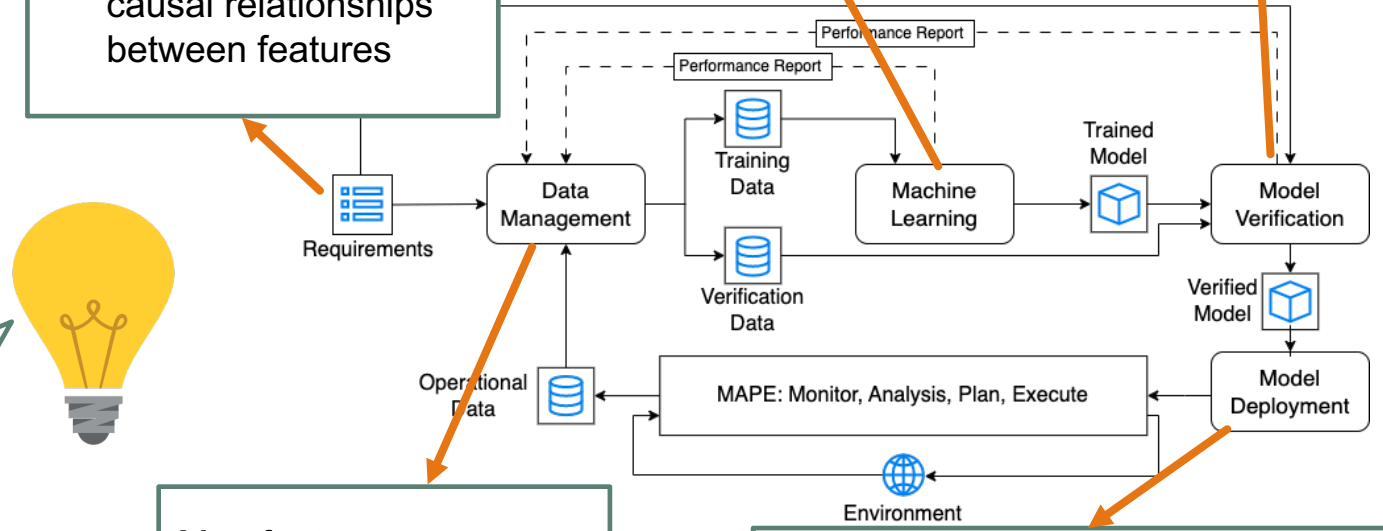
CHALLENGES AND FUTURE DIRECTIONS

With taxonomy,
- map input/output features at each stage
- discover challenges and future research directions
- promote collaboration between SE and other ML development lifecycle experts

- Reduce manual effort
- Handle complex dependencies and causal relationships between features

Interpret ML-learned features

- Use features learned by ML in testing and verification
- Handle complex feature dependencies and interactions
- Enable modular analysis



Map features represented differently in data and requirements

- Identify features
- Monitor ML performance per feature
- Detect unexpected new features and feature interactions during runtime

CONCLUSION

Def.: features are high-level abstractions of desired functionalities, model behaviour, and data

Our vision (feature-based analysis of the ML development lifecycle): use features across all stages of the lifecycle to support software engineers to ...

1) … enable input-output compliance per stage
2) … analyze interactions between different stages
3) … enable modular development of artifacts

Promote collaboration between SE and other experts

THANK YOU

I'm on the job market!
Contact me @
boyue@cs.toronto.edu

Boyue Caroline Hu

UNIVERSITY OF TORONTO