9th UCAAT User Conference on Advanced Automated Testing

Debugging Machine Learning Models

Emmanuel Charleson Dapaah

Prof. Jens Grabowski

UNIVERSITÄT GÖTTINGEN

14/09/2022

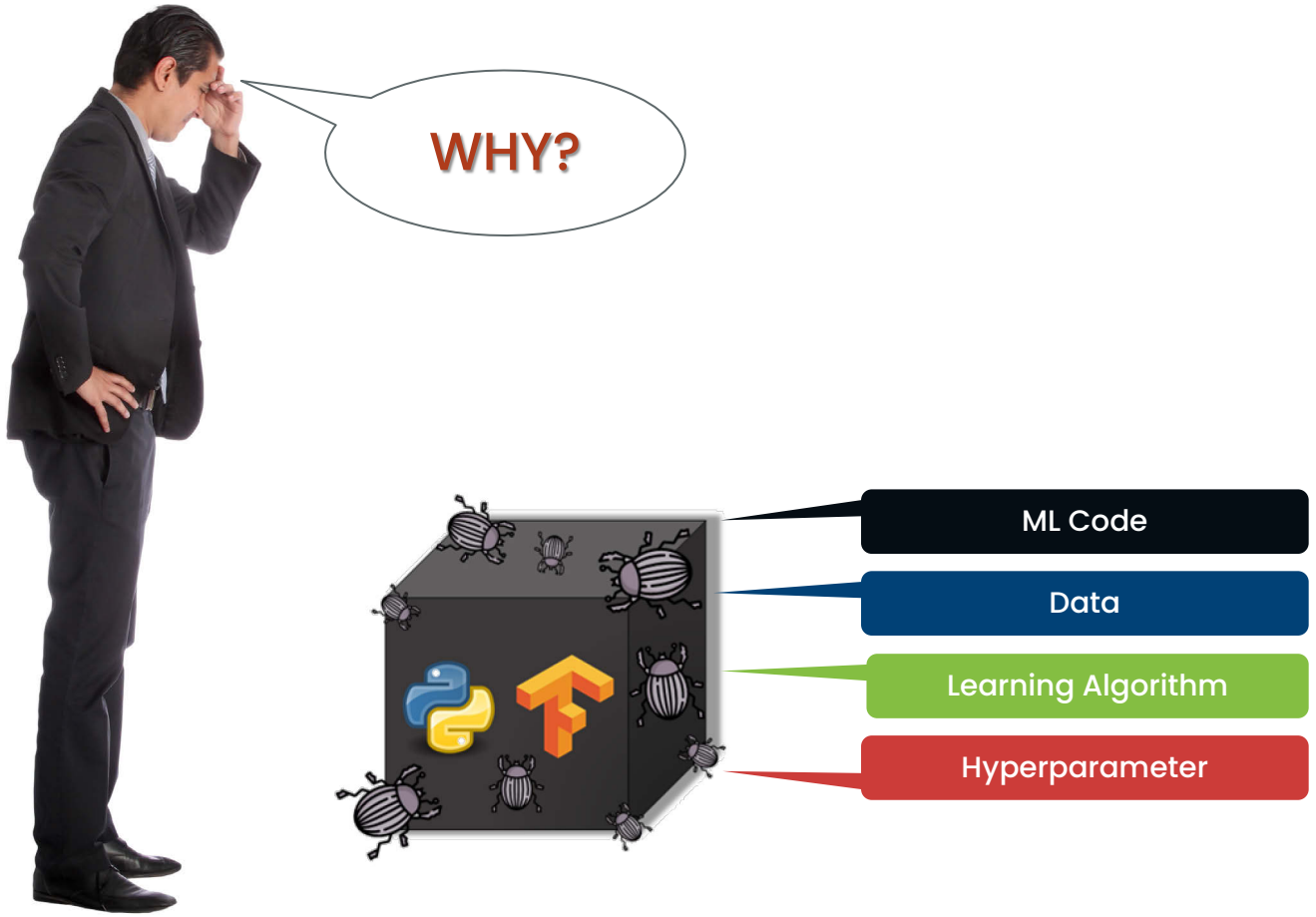
Outline

- Motivation
- Proposed Framework
- Future Works
- Conclusion

Motivation

https://enterrasolutions.com/blog/machine-learning-in-business-time-to-get-excited/

Motivation



Debugging hyperparameter misconfigurations of an ML model using causal reasoning

https://towardsdatascience.com/the-ultimate-guide-to-debugging-your-machine-learning-models-103dc0f9e421

Motivating Example

Misconfiguration

```
model = RandomForestClassifier(bootstrap=True,
                               max_depth=2,
                               max_features='auto',
                               min_samples_leaf=1,
                               min_samples_split=2,
                               n_estimators=30,
                               criterion = 'gini',
                               random_state=42)
```

```
Accuracy:  0.60
```

New User Hyperparameter Grid

```
param_grid = {
    "bootstrap": [True, False],
    "max_depth": np.arange(10, 110, 10),
    "n_estimators": np.arange(100, 500, 50)
    "min_samples_leaf": np.arange(1, 5),
    "min_samples_split": np.arange(2, 4),
    "n_estimators": np.arange(100, 500, 50),
    "criterion": ['gini', 'entropy']
}
```

Can be costly/time-consuming

Debugging

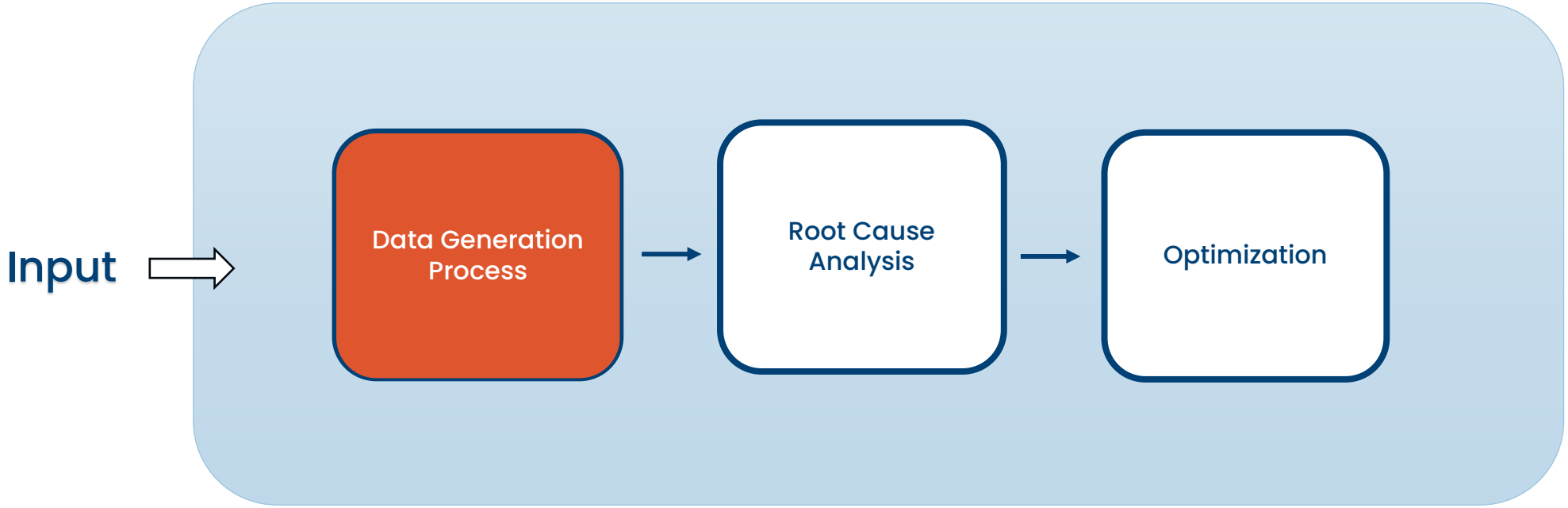
- Root cause
- Reduce the search space
- Interpretability/Explainability

Hyperparameter Optimization

```
model = RandomForestClassifier(bootstrap=False,
                               max_depth=10,
                               max_features='auto',
                               min_samples_leaf=1,
                               min_samples_split=2,
                               n_estimators=400,
                               criterion = 'gini',
                               random_state=42)
```

```
Accuracy:  0.88
```

Proposed Framework (Overview)



Proposed Framework (Data Generation Process)

Input

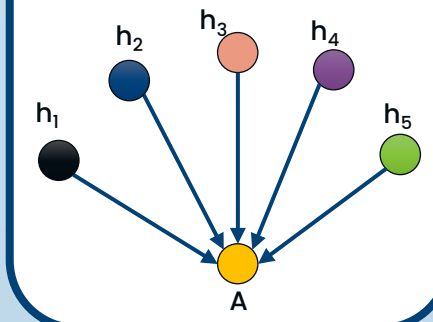
- Dataset
- Learning Algorithm
- User Hyperparameter Grid
- Number of Iterations

Generate Observational data

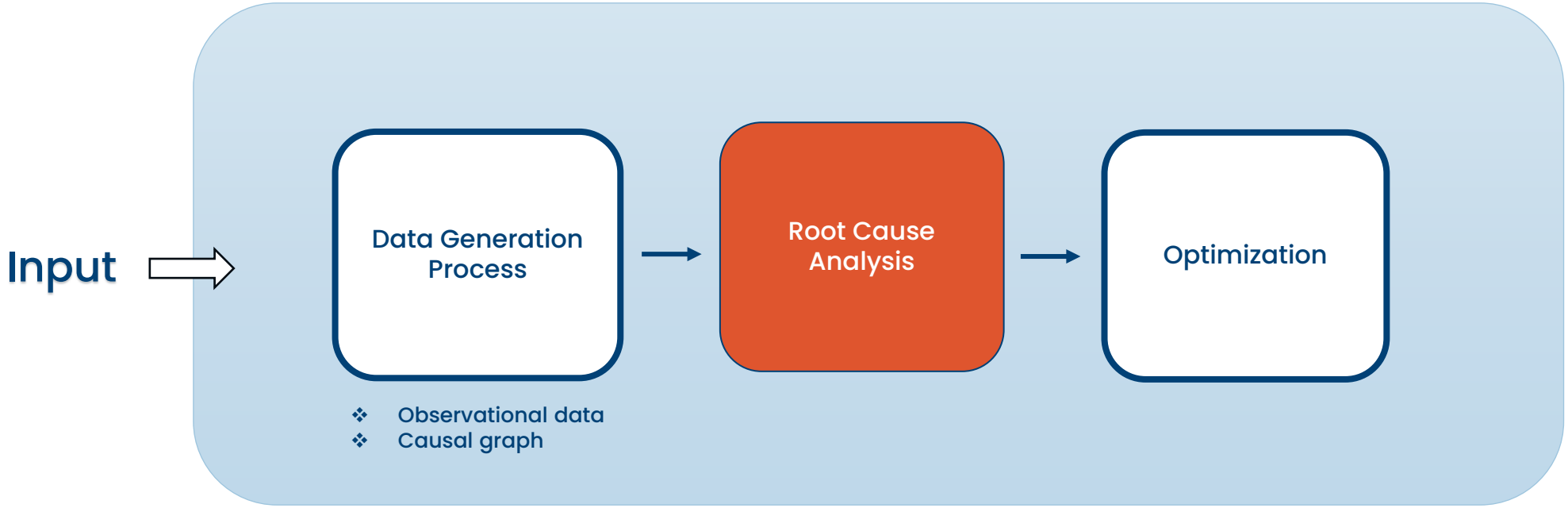
bootstrap	criterion	max_depth	max_features	min_samples_leaf	min_samples_split	n_estimators	accuracy
TRUE	entropy	20	0	9	9	470	84
TRUE	entropy	40	1	10	8	150	62
FALSE	gini	50	0	9	10	180	85
TRUE	entropy	30	0	10	3	250	73

Causality & ML Knowledge

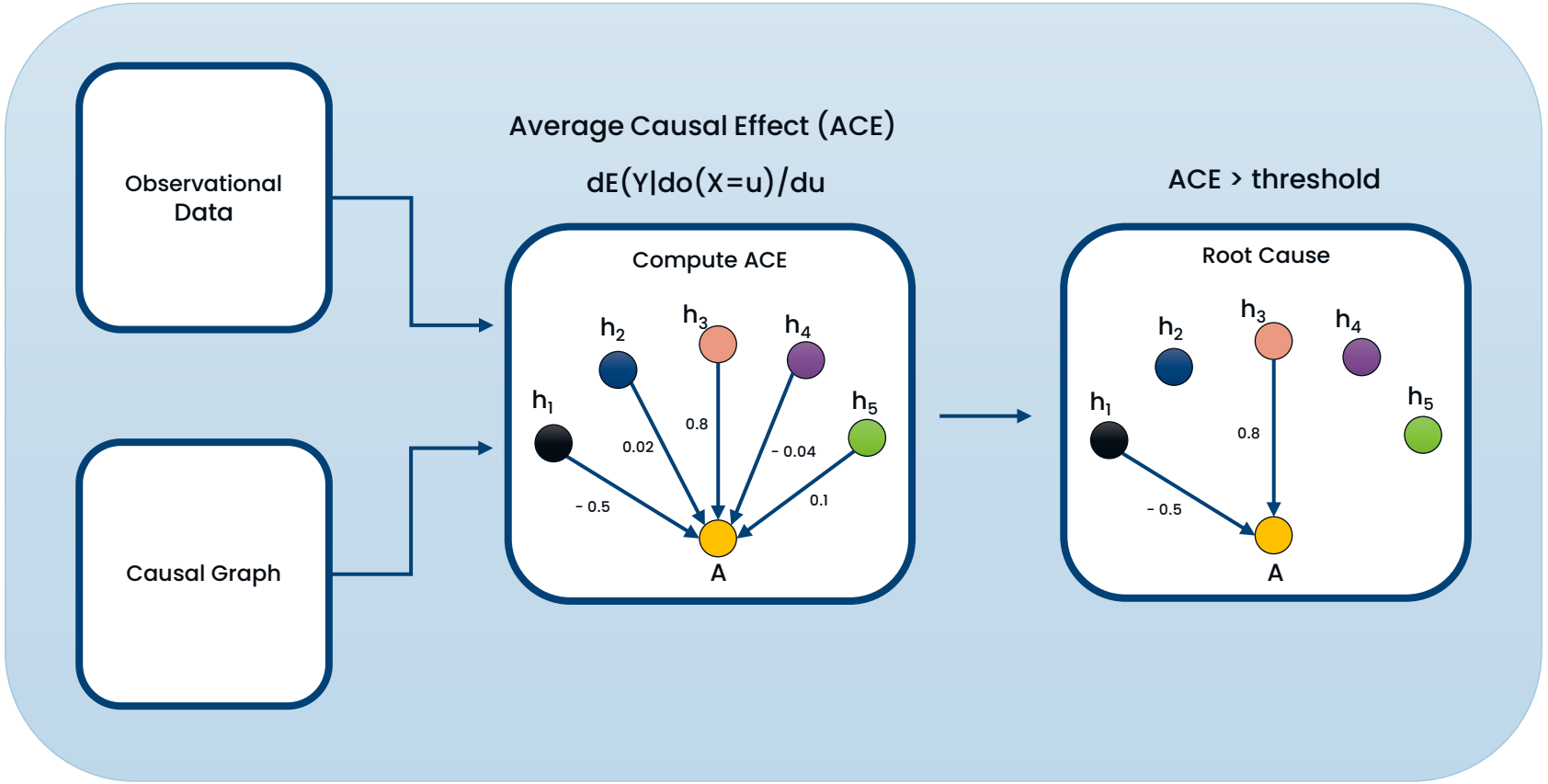
Generate Causal Graph



Proposed Framework



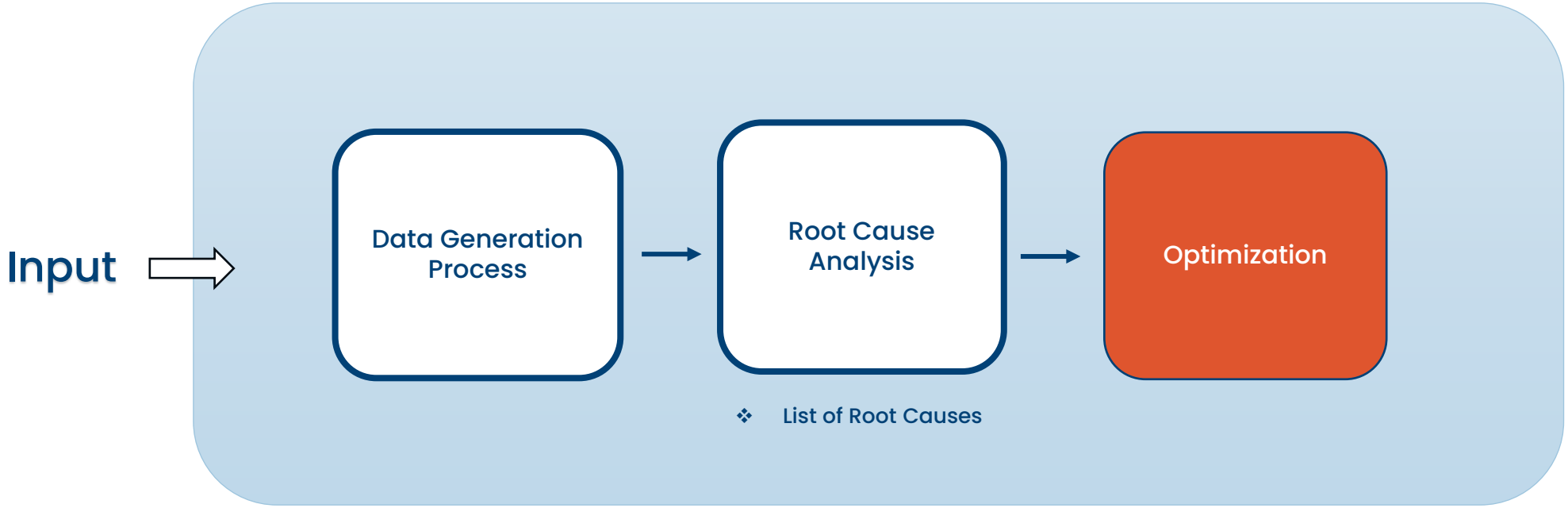
Proposed Framework (Root Cause Analysis)



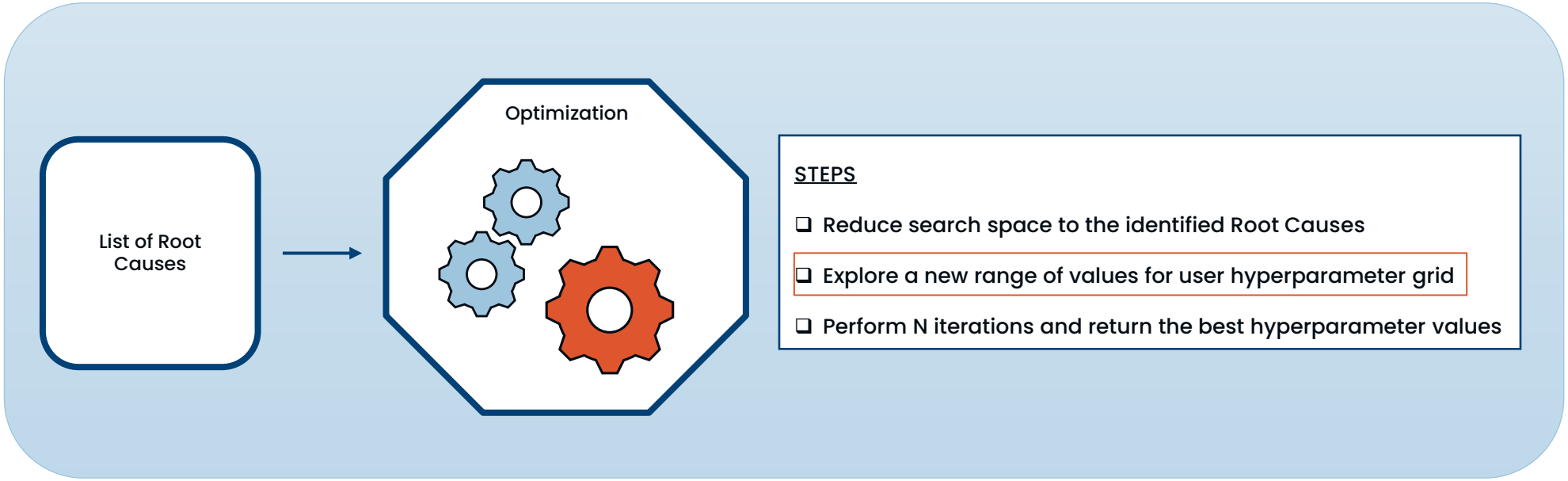
TAKE-AWAY

- ❑ Identify the root causes
- ❑ Infer the tuning direction

Proposed Framework

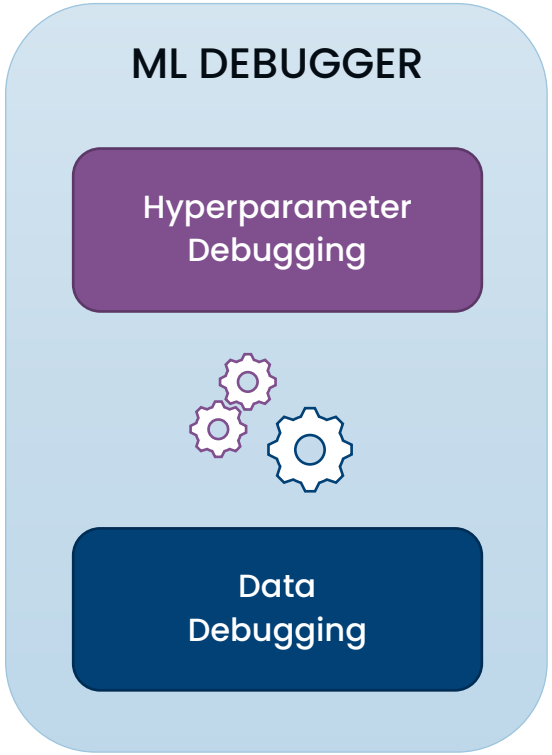


Proposed Framework (Optimization)



Future Works

- Investigate efficient ways of selecting a new range of hyperparameter values
- Evaluate the performance of our framework against existing approaches
- Extend our framework to include data debugging using causal reasoning



Conclusion

- We argued that debugging should be performed before optimization.
- Our framework can help in:
  - Identifying the root cause of a hyperparameter misconfiguration.
  - Reducing the cost involved in hyperparameter optimization and improve on its result.
  - Making the performance of an ML model interpretable/explainable.

Any questions?

(dapaah@cs-goettingen.de)