

Deep Learning Model Complexity: Concepts and Approaches

Xia Hu¹, Lingyang Chu², Jian Pei¹, Jiang Bian³, Weiqing Liu³

¹ Simon Fraser University

² McMaster University

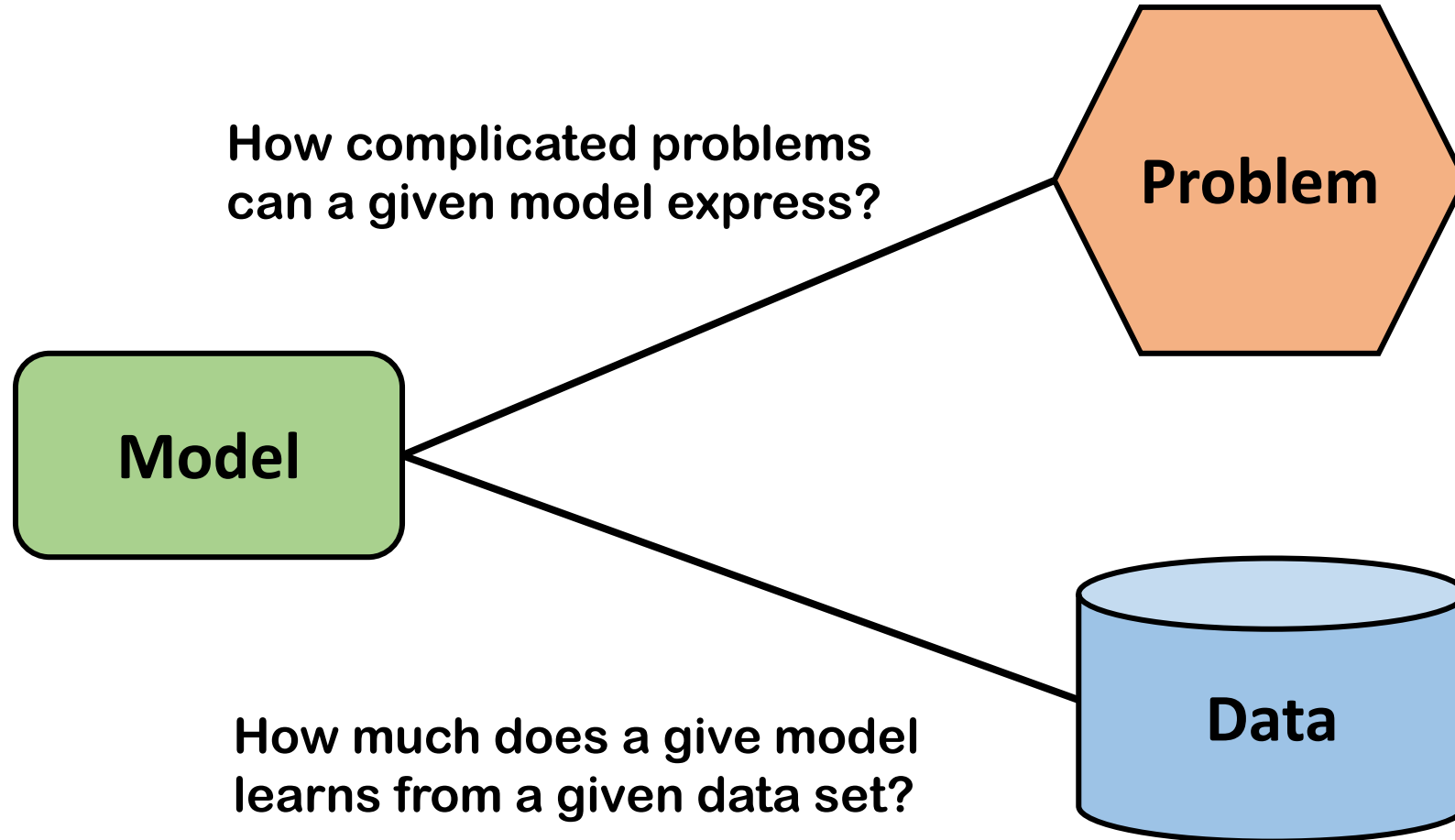
³ Microsoft Research

1

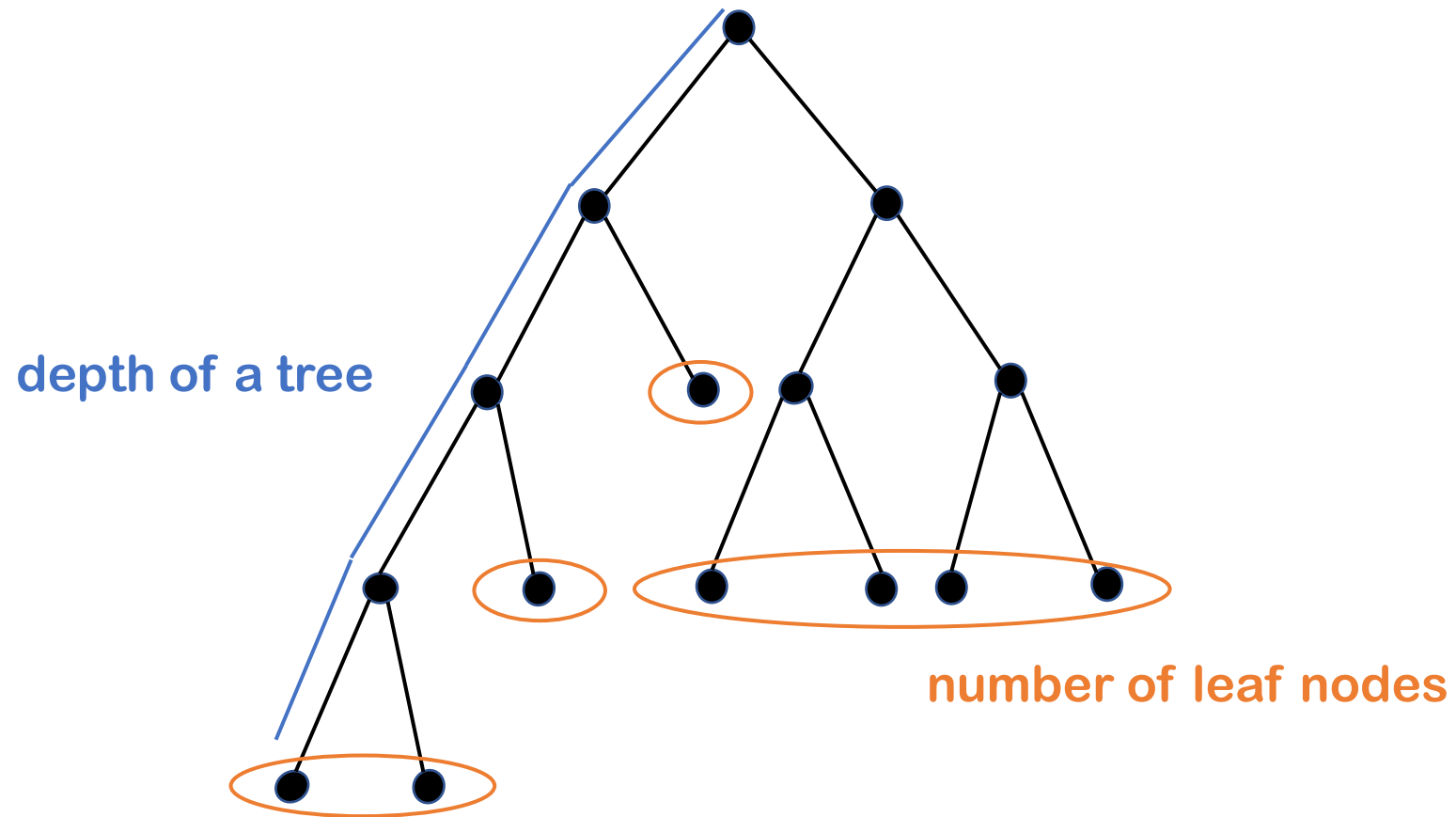
Part I: Introduction

Presenter: Jian Pei

Machine Learning Models and Complexity



Example-1: Model Complexity of Decision Trees



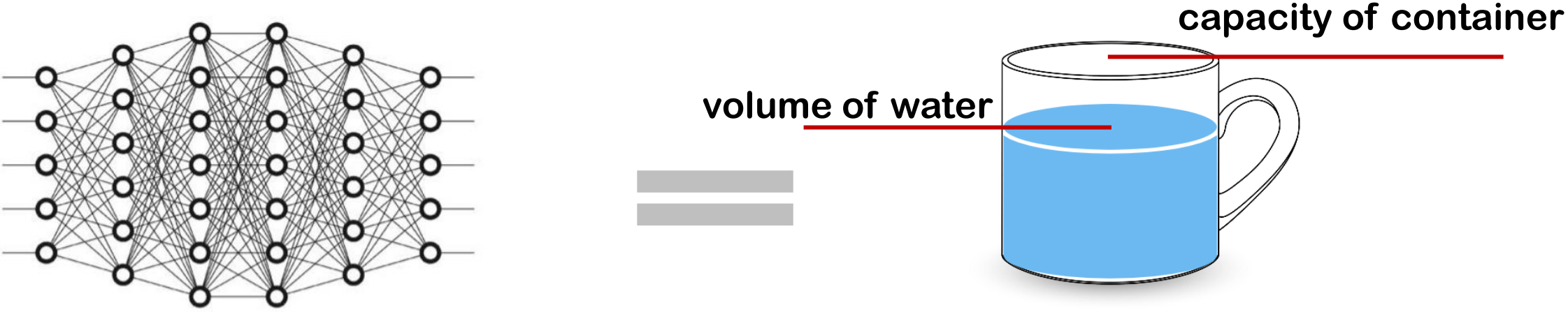
Example-2: Complexity of Logistic Regression

- Vapnik-Chervonenicks theory
- Rademacher complexity
- Fisher Information matrix
- The razor of model
 - A theoretical index of the complexity of a parametric family of models comparing to the true distribution

Deep Learning Models Are Unique

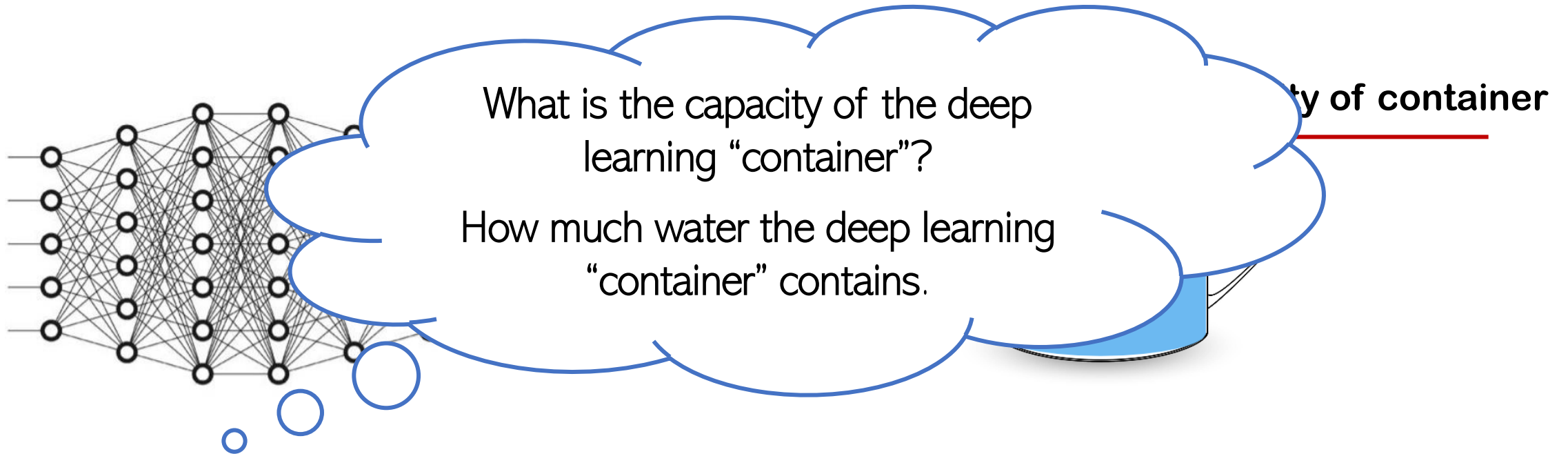
- Large size
- Large number of parameters
- Much higher complexity than traditional machine learning models

What is Deep Learning Model Complexity?



If we informally regard a deep learning model as a “container”.

What is Deep Learning Model Complexity?



If we informally regard a deep learning model as a “container”

Deep Learning Model Complexity

- **Expressive Capacity:** the capacity of deep learning models in approximating complex problems
- **Effective Complexity:** the (practical, usable) complexity of the functions represented by deep models with specific parameterization
- Example
 - Function $f(x) = ax^2 + bx + c$
 - Expressive capacity: unary quadratic
 - Effective complexity when $a = 0$: linear

What is Deep Learning Model Complexity?

- In the hypothesis space H that corresponds to a fixed deep learning model structure
- Effective Complexity
 - the complexity $EMC(h)$ of a specific hypothesis h ($h \in H$)
- Expressive Capacity
 - describes the upper bound of the complexity of any model in H , that is,
 $\sup\{EMC(h):h \in H\}$

Why Deep Learning Model Complexity?

- Understand the capability and limitation of deep learning models
- Investigate many other related fundamental questions
 - Generalization error and generalization capability
 - Overfitting
 - Model optimization and regularization
 - ...

Deep Learning Model Complexity: Factors

- Model framework
 - Model type: FCNN, CNN, RNN, ResNet ...
 - Activation function: Tanh, ReLU ...
- Model size
 - Number of hidden layers, width of each layer, number of filters, number of trainable parameters, ...
- Optimization process
 - The form of objective functions, optimization algorithms, hyperparameters, ...
- Data complexity
 - Data dimensionality, number of class labels, data distribution, ...

Model-Specific vs. Cross-Model Methods

- **Model-specific:** study complexity of a certain type of model
 - Deep neural networks with ReLU activation [Raghu et al., 2017]
 - Deep neural networks with smooth curve activation [Hu et al., 2020]
- **Cross-model:** cover and compare multiple types of models
 - CNNs and RNNs [Khrulkov et al., 2018]

Measure-based vs. Reduction-based Methods

- **Measure-based:** define an appropriate quantitative representation of model complexity

- Number of linear regions of ReLU networks

[Raghu et al., 2017; Hanin and Rolnick, 2019; Hu et al., 2020]

- **Reduction-base:** reduce deep learning models to some known problems and functions

- Connect neural networks to Tensor decompositions [Khrulkov et al., 2018]

Outline of This Tutorial

- Part I: Deep Learning Model Complexity (*Jian Pei*)
- Part II: Expressive Capacity (*Lingyang Chu*)
- Part III: Effective Complexity (*Xia Hu*)
- Part IV: Application Examples (*Jiang Bian*)
- Part V: Conclusion (*Weiqing Liu*)