Optimize the Database Index Structure

Presented by: XingChi Wang

① Opportunity: Why Learned B- tree?

Challenges of Traditional Indexes

- An index is a structure that sorts the values of one or more columns in a database table.
- Most common database index: *B-tree*.
- B-trees know nothing about the distribution of the data.
- B-trees do not take advantage of more common patterns prevalent in datasets.

- Search time for B-Tree index structure is *O(log n)*.
- For a very large *n* this value can be grow extensively.
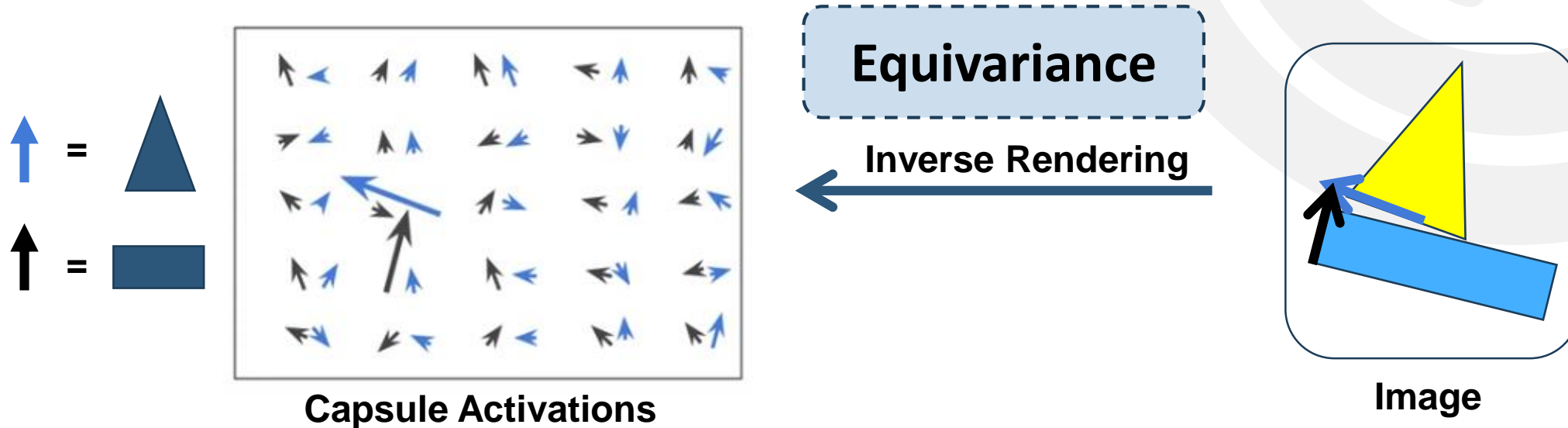- Can we link machine learning with index structure?

Challenges with Search Time
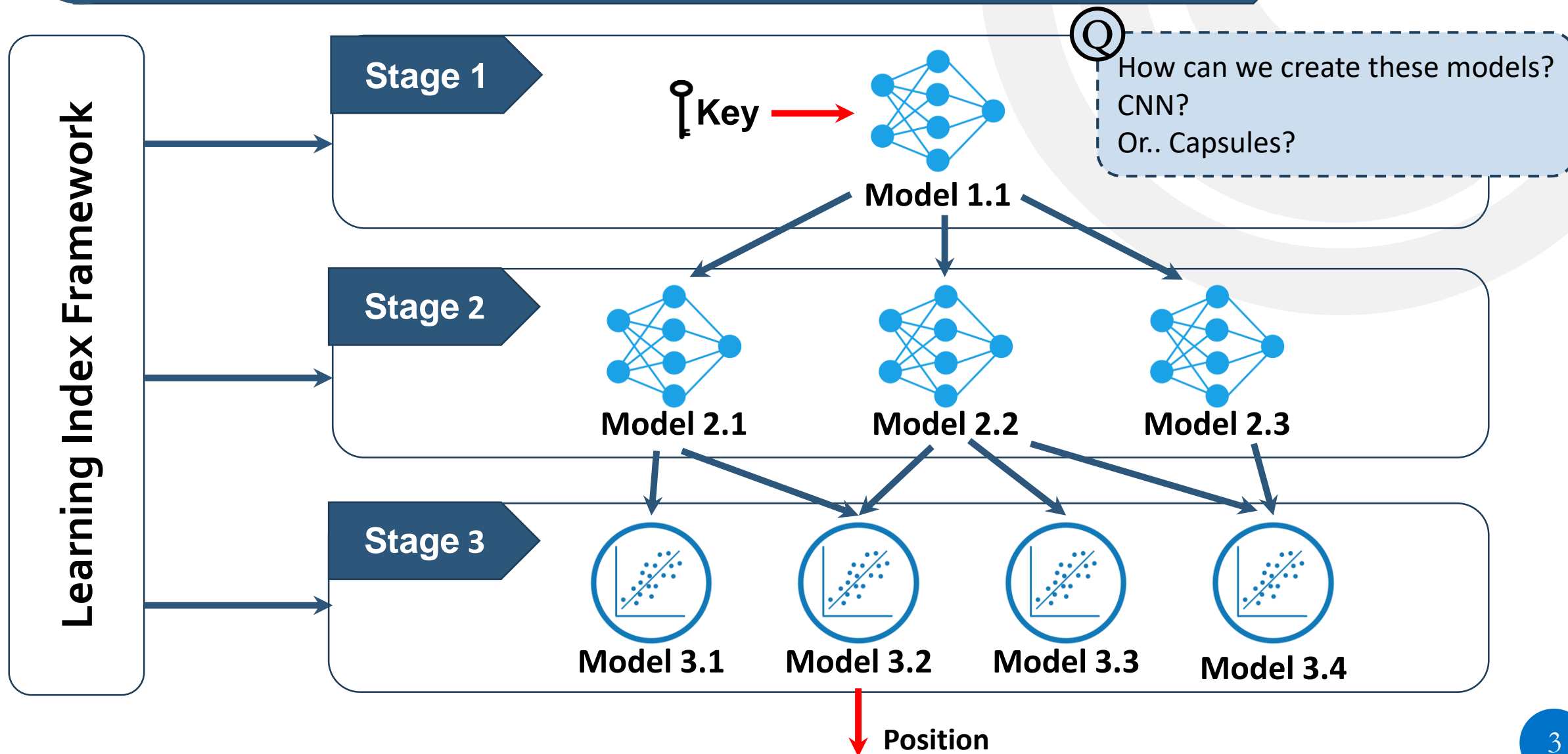
② Solution: Capsule B-Tree?


Capsule Activations

Equivariance

Inverse Rendering

Image

- **Capsule Networks:** are Neural networks that try to do Inverse Graphics.
- Many capsules, each capsule tries to predict the presence and instantiation parameters of an object at a given location.
- Achieved state of the art performance and better results than CNN.

Q How can capsule handle complex data distribution? Composed images?

Strong Agreement

Routing by Agreement

Inverse Rendering

Primary Capsules

Image
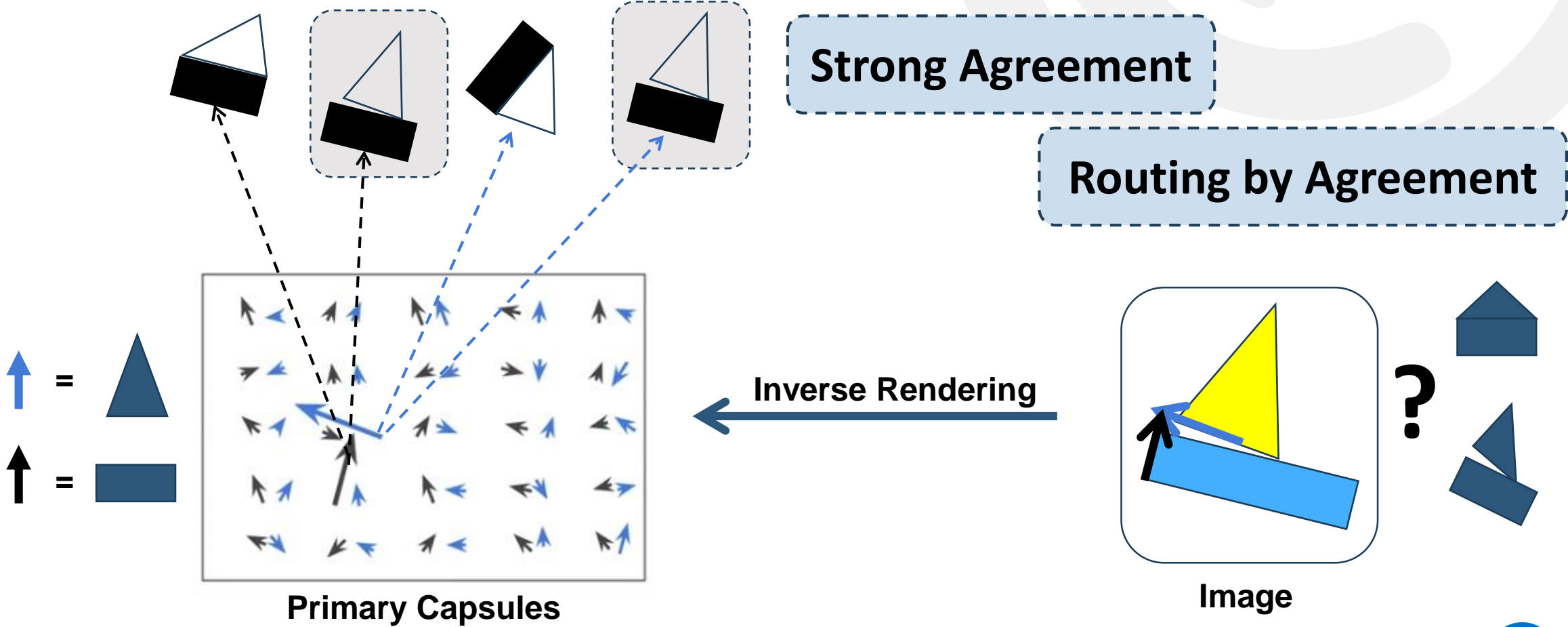
③ Conclusions and Next Steps

Routing by Agreements

- Can handle complex distributions.
- Tends to choose the solution that makes all the capsules achieve accurate predictions for the capsules in the next layers.
- Capsules can be applied to the index problem to outperform CNN.
- Requires less training data.

- Start the implement based upon Capsule Networks approach.
- Collect real-world datasets to start the experiments
- Start experimental study to compare between B- tree and our learned index structure

Next Steps…

UNIVERSITY OF ALBERTA

THANK YOU

✉ nashaata@ualberta.ca

✉ aindrila@ualberta.ca

✉ xingchi@ualberta.ca