Deloitte.

Data Driven Enterprise
Unleashing Machine Learning for Superior Cleansing

Chetna Pandey

SAP Data Migration
Deloitte

Agenda

- The Challenge

- The Solution

- Client Example

- Demo

- About Us

- Questions

The Challenge

Dirty data costs the global economy over $3 trillion every year

Workers spend 60% of their time cleaning, normalising and organising data

Human error accounts for more than 60% of all dirty data

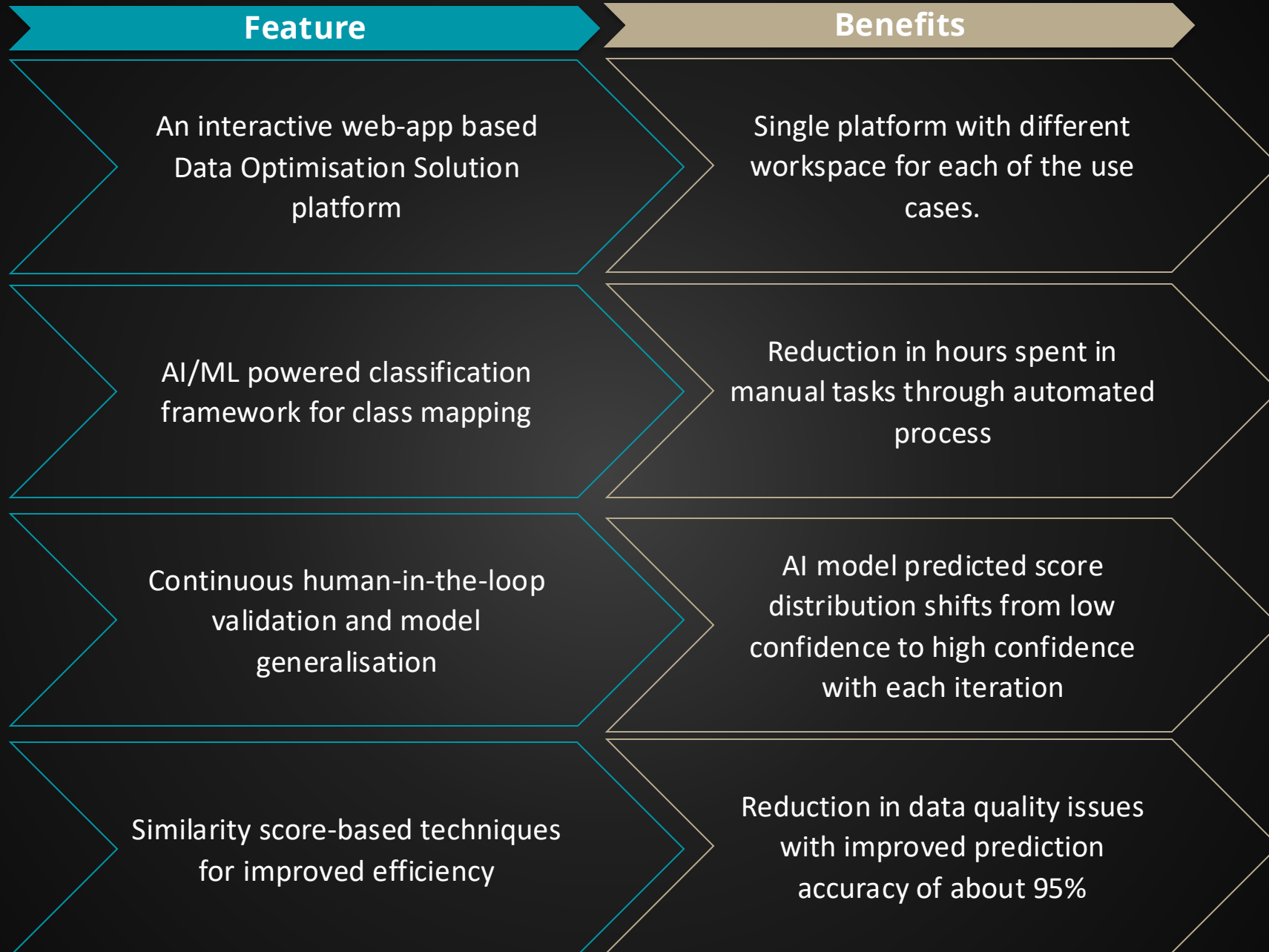**Companies believe, on average, that 26% of their data is corrupt**

Over 57% companies discover dirty data when it's reported

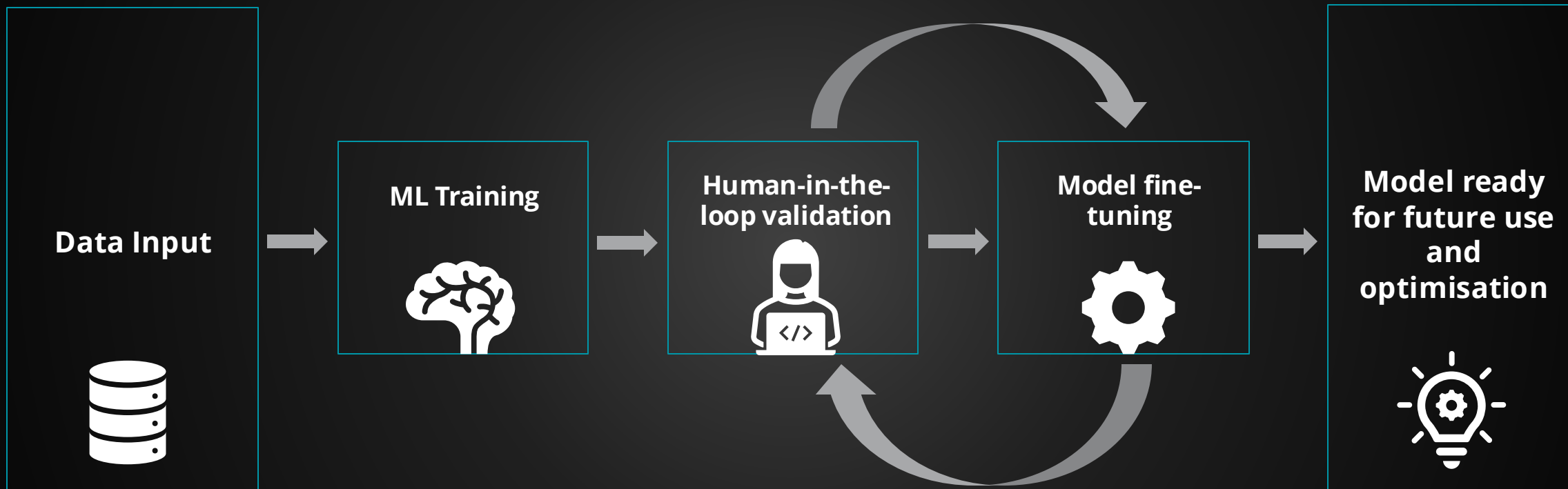**Only 16% executives are confident in the accuracy of their data**

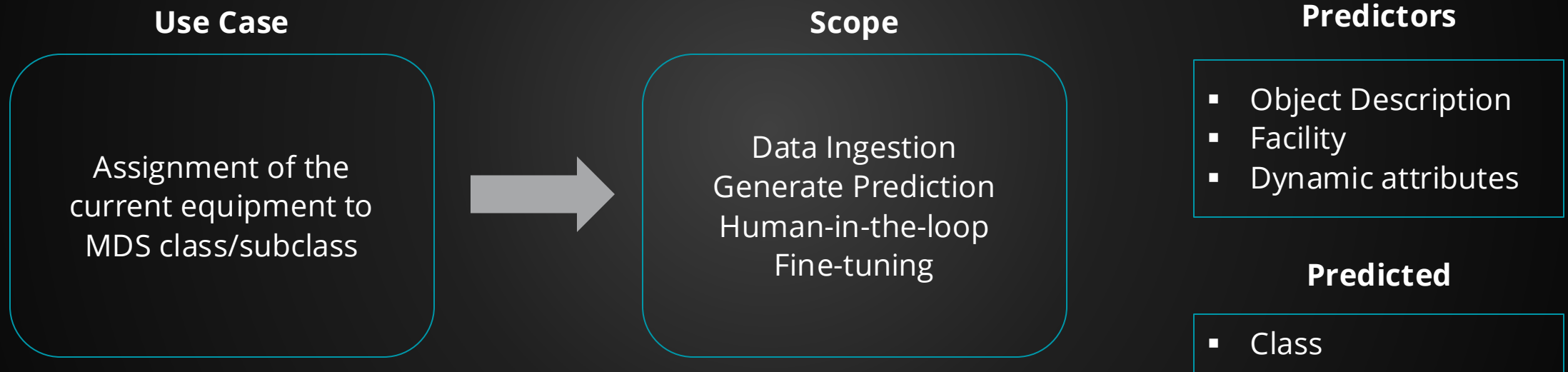# The Solution

Data Optimisation Solution

Feature	Benefits
An interactive web-app based Data Optimisation Solution platform	Single platform with different workspace for each of the use cases.
AI/ML powered classification framework for class mapping	Reduction in hours spent in manual tasks through automated process
Continuous human-in-the-loop validation and model generalisation	AI model predicted score distribution shifts from low confidence to high confidence with each iteration
Similarity score-based techniques for improved efficiency	Reduction in data quality issues with improved prediction accuracy of about 95%

Client Example

Asset Class/Subclass Mapping

Use Case

Assignment of the current equipment to MDS class/subclass

Scope

Data Ingestion
Generate Prediction
Human-in-the-loop
Fine-tuning

Predictors

- Object Description
- Facility
- Dynamic attributes

Predicted

- Class

ML Training

Training Dataset

A	B	C
a1	b1	c1
a2	b2	c2
a4	b4	c4

A	B	C
a1	b1	c1
a2	b2	c2
a3	b3	?
a4	b4	c4
a5	b5	?

Test Dataset

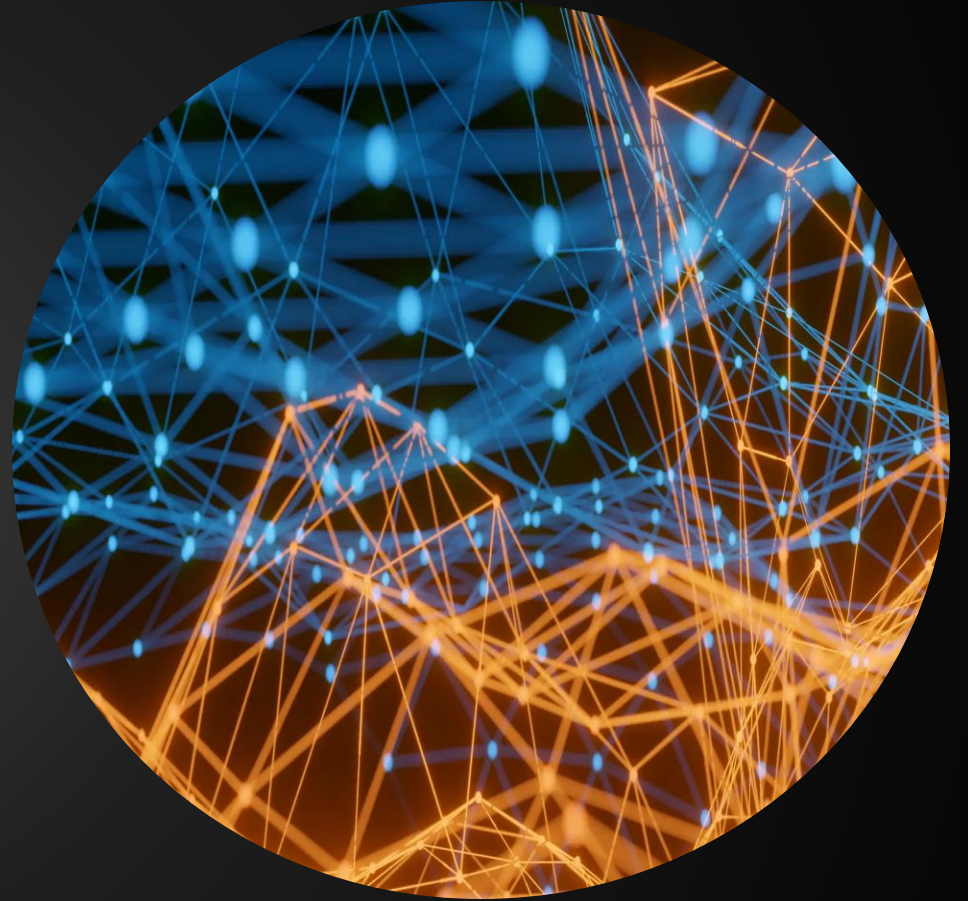| A	B	C
a3	b3	?
a5	b5	?

A

B

Input Layer

Hidden Layer

Output Layer

C
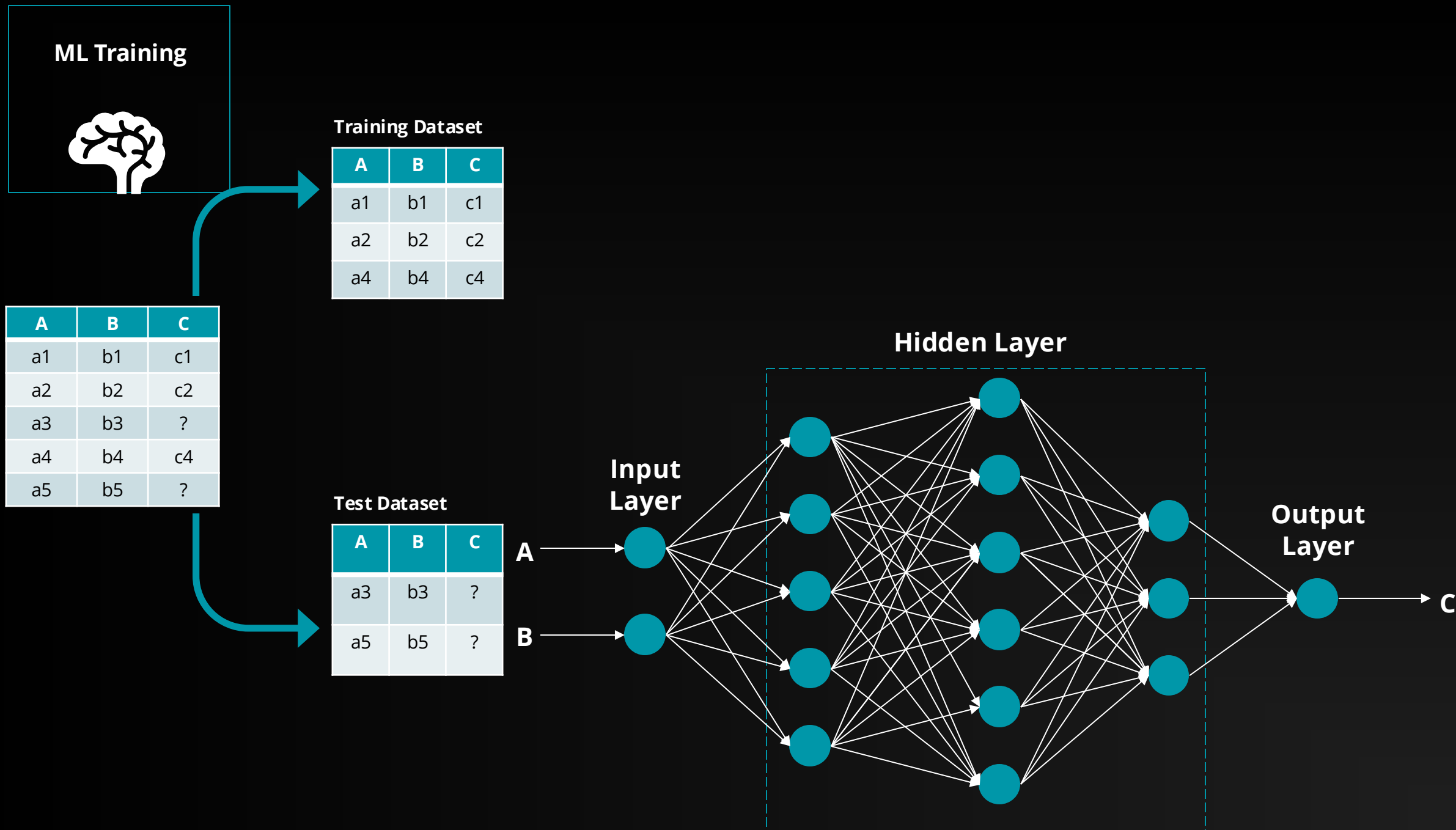
Demo

Questions?

Thank you