



Delzeen
Machhi



Daphne
Kallinderi

PROBLEM OVERVIEW & CURRENT STATE

160+ countries

10,000+ Enterprises

~75 M + connections per day

DATA

28-days telemetry data
85 tables, 2 datacenters
300M+ records, 800+ features

Windows 365 connections via the new UDP protocol

FOCUS

The Windows Cloud Product Group currently performs Root Cause Analysis **manually**. AI can help make diagnostics **smarter** and **faster**.

Involvement of Data Scientists

Product Group

Query 6+ different tables

Check dashboards

100+ hours

Copilot for suggestions

BUSINESS IMPACT

2.4M\$

Cost to resolve connectivity issues in FY25

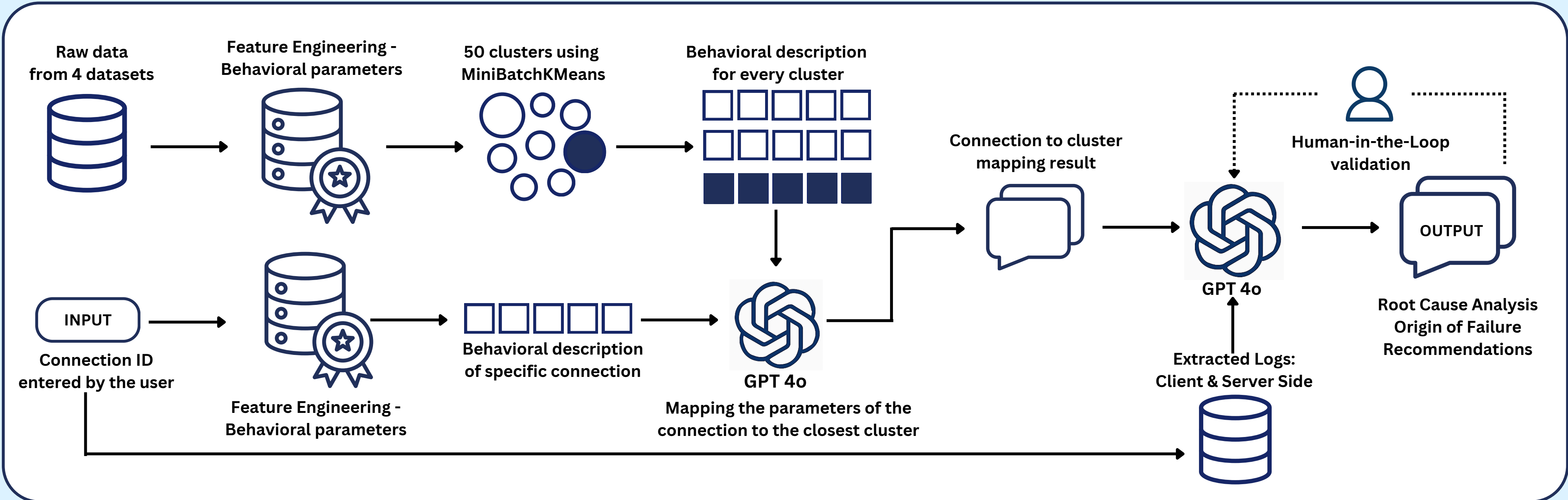
30%

Reduction in time Per case

48%

Reduction in cost Windows Cloud Product Group

METHODOLOGY & WORKFLOW



DATA TRANSFORMATION

EXPLORATORY DATA ANALYSIS: INFLUENCE OF DURATION AND REGION ON ERROR RATE

Different Regions (not displayed for privacy reasons)

Legend: 0-1 mins (blue), 1-30 mins (orange), 30 mins - 8 hrs (green), > 8 hrs (red)

COMPLEXITY OF NETWORK CONFIGURATION

Calculation of distances & Data center congestions

CLIENT GROUPING

~ 100 unique values → 9 client types

ERROR BUCKETING

90+ error codes → 18 logical error buckets

FROM RAW DATA TO MODEL-READY

Large, high-dimensional data - difficult to model → Cleaned, organized & interpretable - ready for analysis

SIMILAR CASES

Learns from similar cases across the globe with multiple metrics combined

UNIFIED VIEW

4 x

2 x

BEHAVIORAL CLUSTERING

FEATURES

- Network metrics
- Session Metadata
- Error Details
- Derived Features

APPROACH

MiniBatchKMeans

10 Features
k = 50

Bandwidth	RoundTrip Time	TimeSlot	Client Region	Client Type
Very High	Bad	Morning	New York	Windows

Clusters trained on **successful** and **failed** connections

Cohesive Study of Metrics affecting connection reliability

CHALLENGES

- Domain Complexity
- Scalability Constraints
- Limited Data Retention
- Dynamic Network Configurations
- Error Code Imbalance
- Lack of Relay Location Data