

# AI-Driven Entity Intelligence & Risk Analysis





# Overview

- ▶ Introduction
- ▶ Approach
- ▶ Risk Parameters
- ▶ Risk Trends for the Trained data
- ▶ Results





# Introduction

The goal of this challenge is to build an AI-driven system that:

- Extracts entity names from unstructured and structured transaction data.
- Enriches the extracted names with publicly available data (e.g., company registries, online sources, financial news, regulatory filings, and legal databases).
- Identifies potential fraudulent or high-risk entities through anomaly detection.
- Classifies entities into categories (corporation, non-profit, shell company, government agency, etc.).
- Assigns a risk score based on entity attributes, and associated networks (business/ sectors associated with the entities)
- Provides supporting evidence and confidence scores to assist analysts in decision- making.



# Approach



## **Training Dataset Preparation:**

- We start with a dataset containing 100 entries, each with 10 parameters used to calculate the risk score.
- Based on the ratings of these parameters, we assign weights to each of the 10 parameters.
- The risk score is then computed by rounding off the weighted values.
- This dataset, including the calculated risk scores, serves as our TRAIN DATASET for model training.

## **Model Training**

- The TRAIN DATASET is used as input to train the machine learning model.
- The model learns to predict risk scores and classify entity types based on the given parameter ratings and assigned weights.

## **Testing and Prediction:**

- A new dataset (TEST DATASET) is provided, containing only parameter ratings for different transactions.
- The trained model processes this dataset to predict risk scores and classify entity types.

## **Automated Output Generation:**

- The model generates an output file in Excel format.
- This file contains the predicted risk scores and entity types for the given TEST DATASET, ensuring an automated and structured risk assessment.

# Risk Parameters

01

Jurisdiction Risk – Risk based on the regulatory environment.

02

Shell Company Detection – Identifies potential shell entities.

03

Beneficial Ownership Transparency – Evaluates ownership structure transparency.

04

Sanctions & Watchlist Screening – Checks involvement with restricted entities.

05

Round-Number Transactions – Identifies transactions with rounded amounts, often linked to fraud.

06

Politically Exposed Persons (PEP) Involvement – Assesses if high-profile individuals are involved.

07

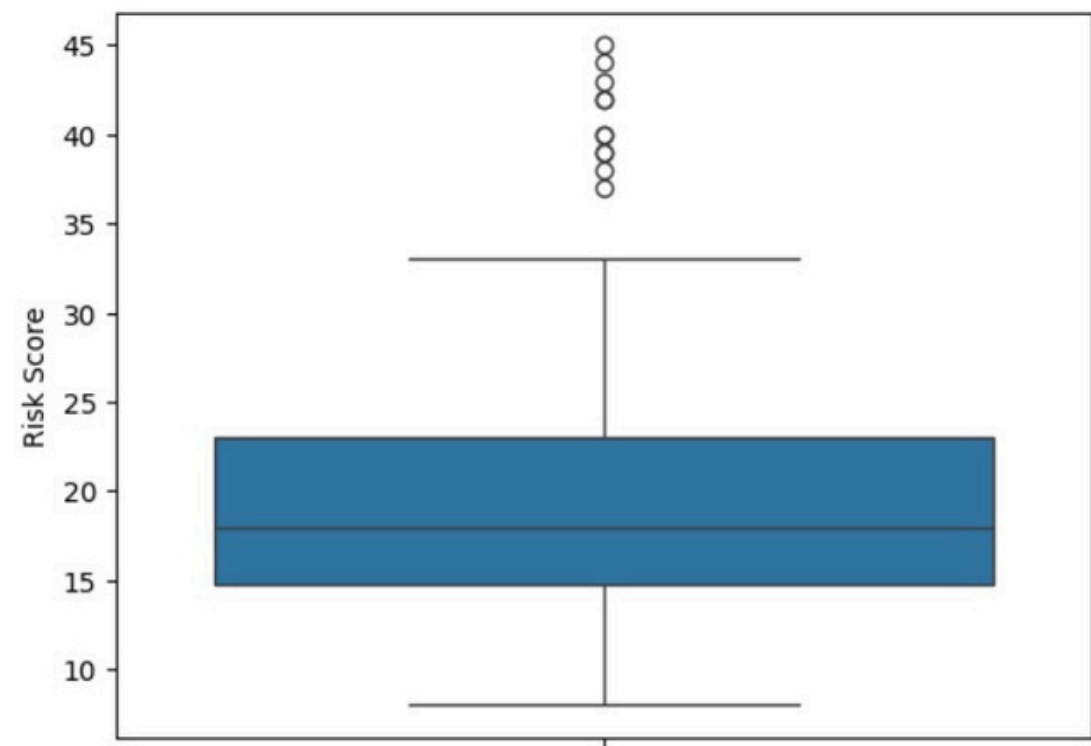
Unusual Transaction Timing – Flags transactions with irregular patterns.

08

Industry-Specific Risk Factors – Considers risks specific to industries.



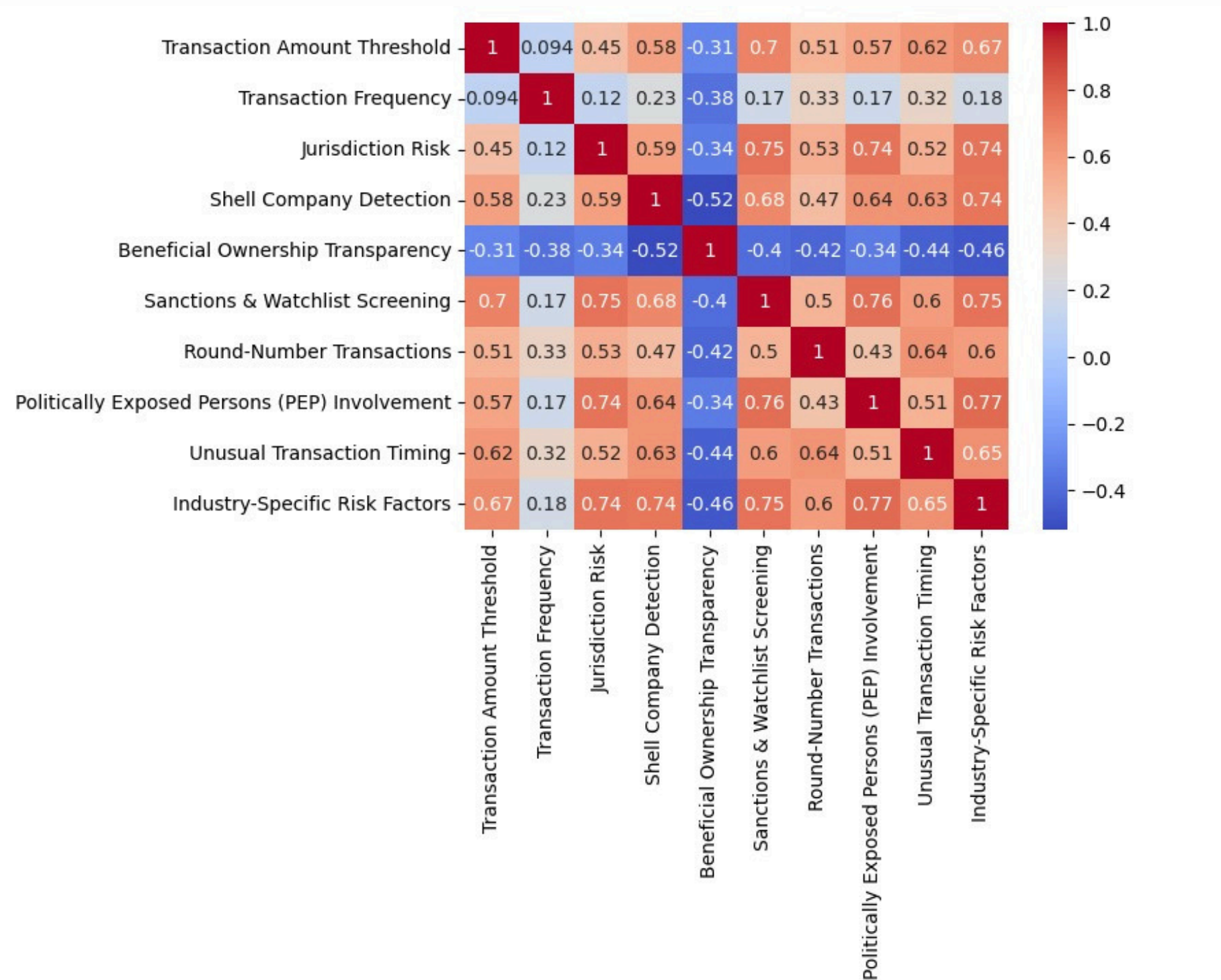
# Risk Trends for the Trained data



Model Accuracy

|              |  |        |          |         |  |
|--------------|--|--------|----------|---------|--|
| [↕]          | Mean Absolute Error (Risk Score): 1.2010000000000003 |        |          |         |  |
|              | Classification Report (Risk Category):               |        |          |         |  |
|              | precision  | recall | f1-score | support |  |
| High         | 1.00   | 1.00   | 1.00     | 4       |  |
| Low          | 0.50   | 0.25   | 0.33     | 4       |  |
| Medium       | 0.79   | 0.92   | 0.85     | 12      |  |
| accuracy     |  |        | 0.80     | 20      |  |
| macro avg    | 0.76   | 0.72   | 0.73     | 20      |  |
| weighted avg | 0.77   | 0.80   | 0.77     | 20      |  |

Correlation Analysis



# Result

predicted\_transactions.csv X

1 to 10 of 13 entries Filter

| Predicted Risk Score | Predicted Risk Category | Transaction Amount Threshold Risk Level | Transaction Frequency Risk Level | Jurisdiction Risk Risk Level | Shell Company Detection Risk Level | Beneficial Ownership |
|----------------------|-------------------------|---|----------------------------------|------------------------------|------------------------------------|----------------------|
| 20                   | Low                     | Medium                                  | Medium                           | Low                          | Medium                             | Low                  |
| 20                   | Medium                  | Low                                     | High                             | Medium                       | Medium                             | Low                  |
| 15                   | Medium                  | Low                                     | High                             | Low                          | High                               | Low                  |
| 25                   | Medium                  | Medium                                  | Medium                           | Low                          | Medium                             | Medium               |
| 19                   | Medium                  | Medium                                  | Low                              | Medium                       | Medium                             | High                 |
| 17                   | Medium                  | Medium                                  | High                             | Medium                       | High                               | Medium               |
| 21                   | Medium                  | Low                                     | High                             | High                         | Medium                             | Low                  |
| 14                   | Medium                  | Medium                                  | High                             | Low                          | Low                                | High                 |
| 17                   | Medium                  | Low                                     | High                             | High                         | Medium                             | Low                  |
| 19                   | Medium                  | High                                    | Low                              | Low                          | Medium                             | Low                  |

Show 10 per page 1 2

Output generated includes the risk scores based on the predictions done by our ML model through the TRAIN DATASET that we've used to understand the patterns of how risk score and entity types are calculated. A new excel file is automatically generated as an output.

# THANK YOU!

