

# PITCH DECK

Detection of Fraudulent Bank Transactions

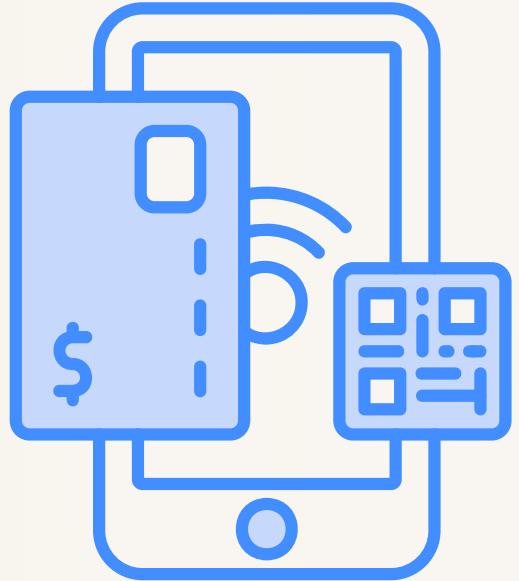
George Shameih, BOUCHABCHOUB Badr, ROUGEMONT Luca,  
BEN RHOUMA Seifeddine, CHARYEV Serdar, BERNHART Eric

# TABLE OF CONTENTS



- Introduction & Problem Statement
- Project Objective
- Dataset Overview
- Methodology & Approach
- Results & Model Performance
- Dashboard & Visualizations
- Conclusion

# THE PROBLEM



**Fraud is the act of deceiving a financial institution or user to gain unauthorized monetary benefit.**

Global card fraud losses exceeded \$12.5 billion in 2024  
(Federal trade commission)

Bank fraud continues to grow year after year.  
A 24% rise between 2023 and 2024.



# PROJECT OBJECTIVE



Develop a Machine Learning model capable of predicting whether a banking transaction is fraudulent (1) or legitimate (0) using historical transaction, user, and card data.

# DATASET OVERVIEW

**210 000 transactions**

File Name	Description
<b>transactions_train.csv</b>	<b>Training data containing transaction features (amounts, dates, merchants, etc.)</b>
<b>train_fraud_labels.json</b>	<b>Labels for training transactions (1 = fraud, 0 = non-fraud)</b>
<b>cards_data.csv</b>	<b>Information about payment cards (e.g., card type, issuer)</b>
<b>users_data.csv</b>	<b>User profile and demographic data</b>
<b>mcc_codes.json</b>	<b>Merchant Category Codes and descriptions</b>
<b>evaluation_features.csv</b>	<b>Evaluation transactions (no labels, used only for prediction)</b>

# METHODOLOGY & APPROACH



Data preparation

type conversion, table joins

Models tested with auto AI

Logistic Regression,  
Random Forest, XGBoost...

Evaluation metrics

ROC-AUC,  
precision, recall,  
F1-score

Validation strategy

ensuring generalization to new (unseen) data

Projets / fraude / transactions

Étapes (5)

Utiliser un modèle de code pour ajouter une étape

Données Profil Visualisations

Source de données : transactions\_train.csv

1. Convertir le type de colonne : Conversion automatique d'une ou de plusieurs colonnes en types de données déduits. Les chaînes qui sont converties au format décimal utilisent un point (.) pour le symbole décimal.

2. Convertir le type de colonne : Types de données convertis manuellement pour 4 colonnes.

3. Jointure : Source de données secondaire cards\_data.csv

L'opération a effectué une jointure de type full sur les données de cards\_data.csv à partir des colonnes card\_id\_id

Nouvelle étape + Configurer ▾ Affichage : 10000 ligne(s), 38 colonne(s) Ensemble de données échantilloné : 8978 ligne(s), 12 colonne(s) ▾

	transaction_id	date	client_id_x	card_id	amount
1	19765990	2017-07-07 09:53:...	1581	2519	\$12.35
2	22160255	2018-11-26 17:44:...	1862	4049	\$58.29
3	17566794	2016-03-26 12:42:...	1967	3367	\$11.03
4	17318690	2016-02-01 08:30:...	921	3457	\$85.74
5	20994060	2018-03-24 14:42:...	456	2800	\$13.43
6	20501200	2017-12-09 22:26:...	0	4639	\$53.12
7	18454176	2016-10-01 21:40:...	408	4960	\$52.86
8	21998064	2018-10-23 12:50:...	1432	4929	\$35.33
9	18977227	2017-01-21 09:30:...	177	2681	\$188.77
10	17244867	2016-01-16 13:40:...	1895	1090	\$27.40
11	17266754	2016-01-21 06:30:...	846	3443	\$4.43
12	20085114	2017-09-12 20:08:...	1904	1004	\$54.00
13	20446347	2017-11-28 11:58:...	696	39	\$100.41
14	20741827	2018-01-29 22:37:...	1439	6094	\$31.98
15	19788807	2017-07-12 08:07:...	762	1080	\$63.54
16	17292719	2016-01-26 15:30:...	1358	3426	\$26.96

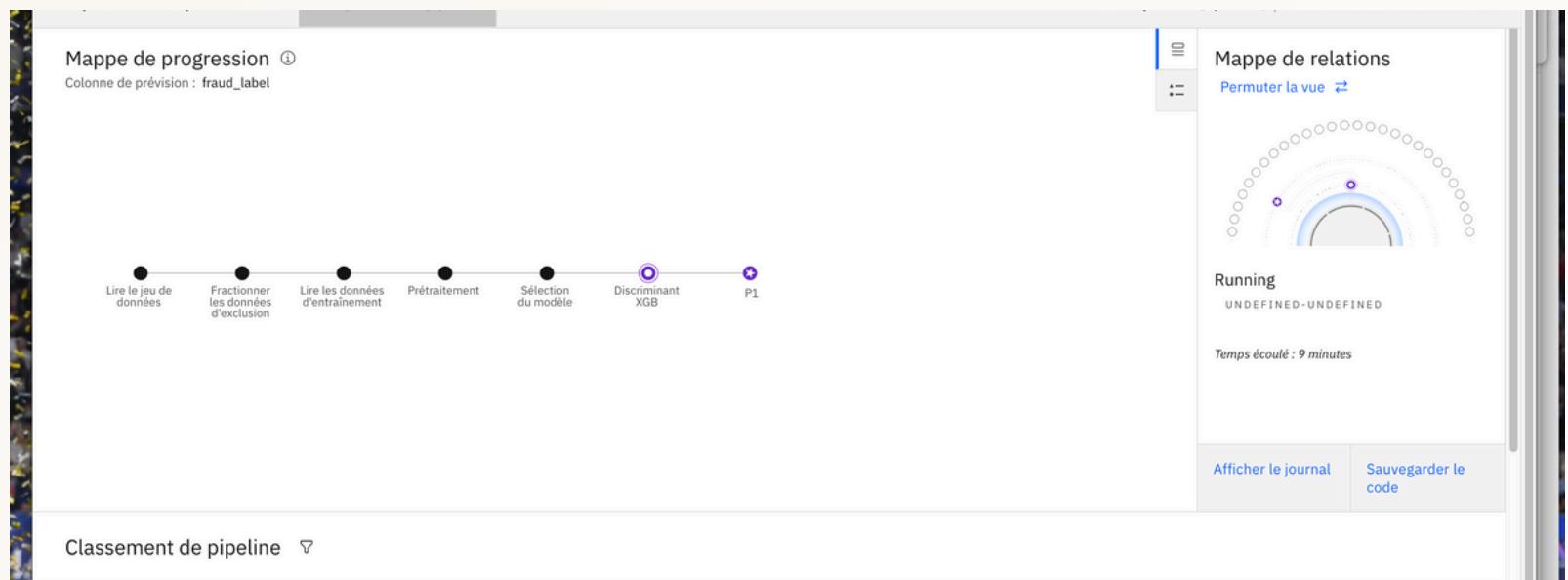
# RESULTS & MODEL PERFORMANCE



XGboost



Accuracy 0.99  
F1 0.71



# DASHBOARD & VISUALIZATIONS

## Fraud Detection Dashboard

ESILV Hackathon 2025 - Track Finance

Powered by IBM Watson X

07/11/2025 10:45:51

Dashboard Data Upload Preprocessing Model Training Models Prediction Visualization

Welcome to Fraud Detection Dashboard!  
Système de détection de fraude bancaire utilisant IBM Watson X.ai - ESILV Hackathon 2025

**210000**  
Total Transactions  
Dataset complet Watson X

**0.15%**  
Taux de Fraude  
Cold-start challenge

**0**  
Modèles Entraînés  
AutoAI & Custom

**93.7**  
Meilleure Précision  
ROC-AUC Score

**Pipeline Progress**

Data Upload: Complete (100%)

Preprocessing & Feature Engineering: Complete

**System Status**

Watson X Connection: Connecting to Watson X.ai...

Dataset Status

# **THE END**

Thank Your For Listening