

# SunSafe

“Shining a Light on Energy Integrity”



Presented by Team  
Sunflower:  
**CTO - Annette Manntschke**  
**COO - Tim Weiand**  
**CHO - Utku Ozkan**



**Website**  
[www.SunSafe.com](http://www.SunSafe.com)

“Shining a Light on Energy Integrity”

# SunSafe Sentinel

SunSafe specializes in cutting-edge AI-powered solutions for fraud detection. We help modern energy providers ensure security and efficiency.

2



**Website**  
[www.SunSafe.com](http://www.SunSafe.com)

# Mission Statement

## Fraud detection

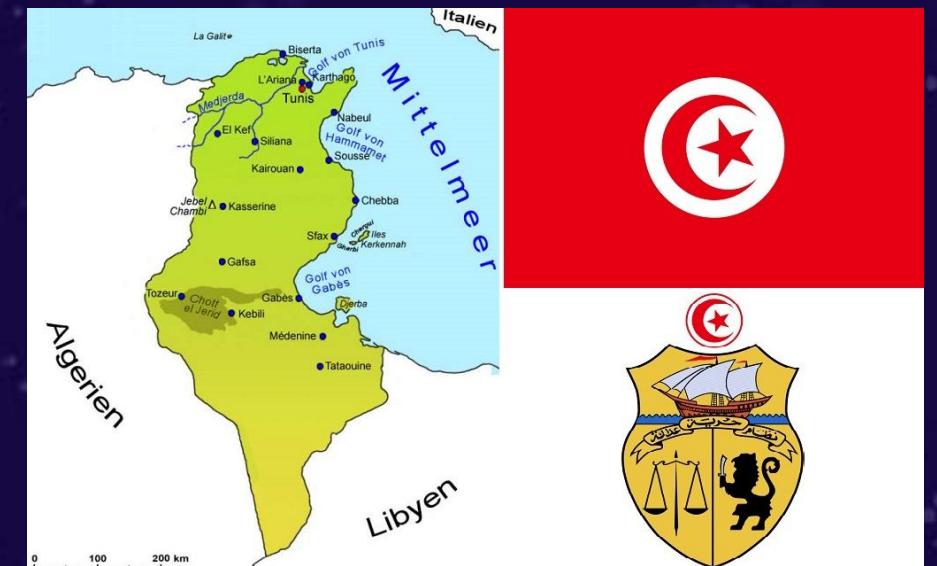
- Meter manipulations cost energy providers a lot of money: STEG (Société Tunisienne de l'Électricité et du Gaz) annually loses 300k Euros due to manipulations of meters.
- We help ensuring fair and transparent energy consumption for all.
- Leveraging advanced machine learning and data-driven insights is our specialty.
- We empower utility companies to protect their revenue, enhance operational efficiency, and build a more sustainable energy future.

## Custom models and updates

Revolutionary Two-stage AI model:

1. maximize fraud case detection using XGBoost
2. refinement of accuracy with a secondary layer

We provide our customers like Steg with an adaptable solution that flags fraud and categorizes risk levels



## STEG client data EDA

# Preprocessing

unprocessed data



Clean, aggregated data  
from 2005 to 2018

- We refined the STEG data by cleaning inconsistencies and aggregated billing records per customer to construct a well structured dataset for our analysis.
- Features for gas and electricity bills included: e.g. region, district, STEG consumption category of each client
- We understood the class imbalance in the client data and later addressed it for accurate and reliable fraud detection

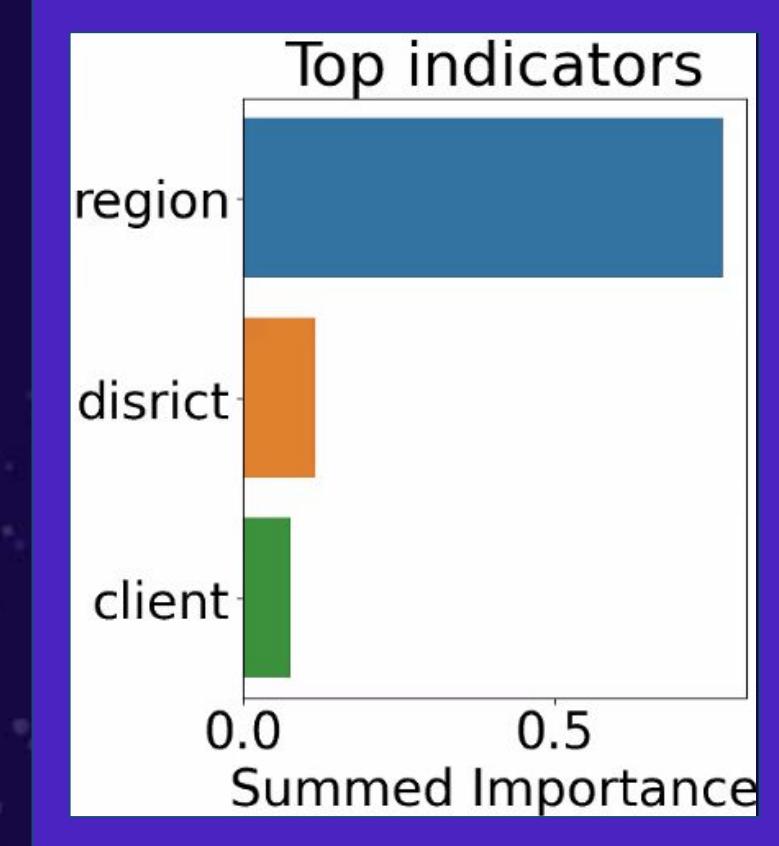
94%

Non-fraud

6%

Fraud

# Layered ML model



01

## Preprocessing

- **Imbalance:** assigned more weight to fraud cases
- **Imputing:** missing values
- **Encoding:** categorical features

02

## XGBOOST layer

- fast
- no scaling
- good generalization
- rank feature impact
- optimized for finding all fraud cases

03

## 2nd Layer

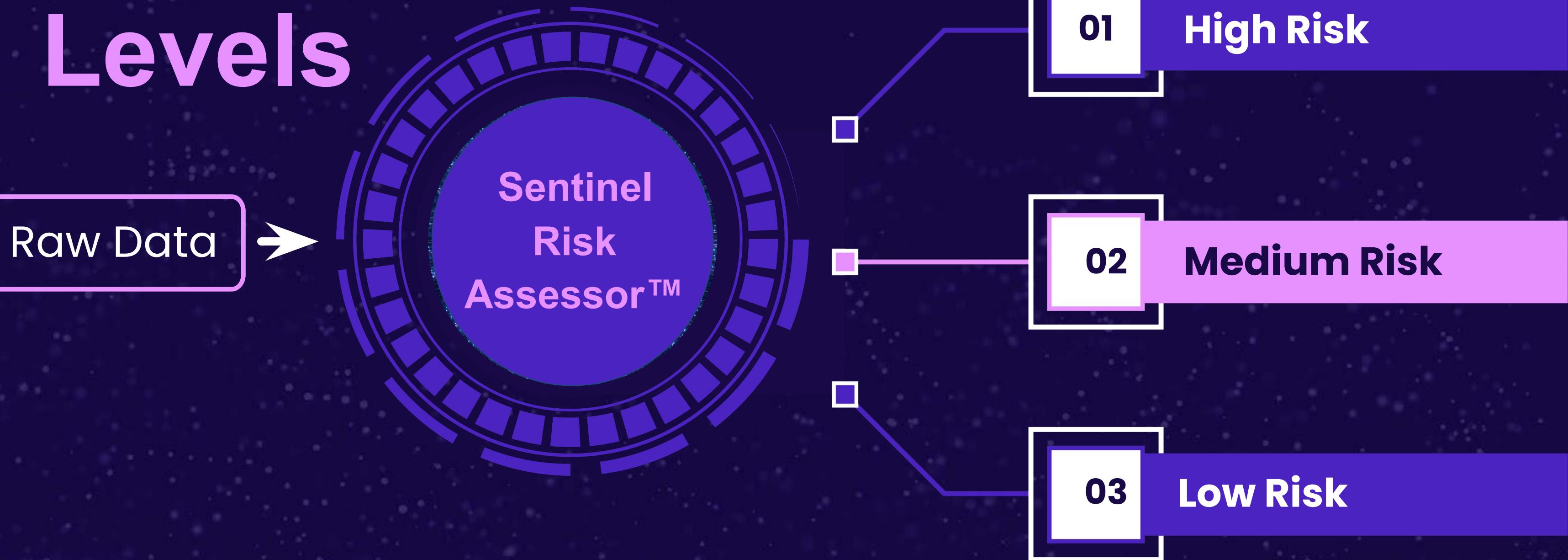
- trained on output of first stage
- optimized for eliminating false positives
- including categorization into risk levels with adjustable thresholds

What we provide for STEG:  
Customized Solution!

Three level system:  
provides flexible human  
verification workload

# Categorizing Risk

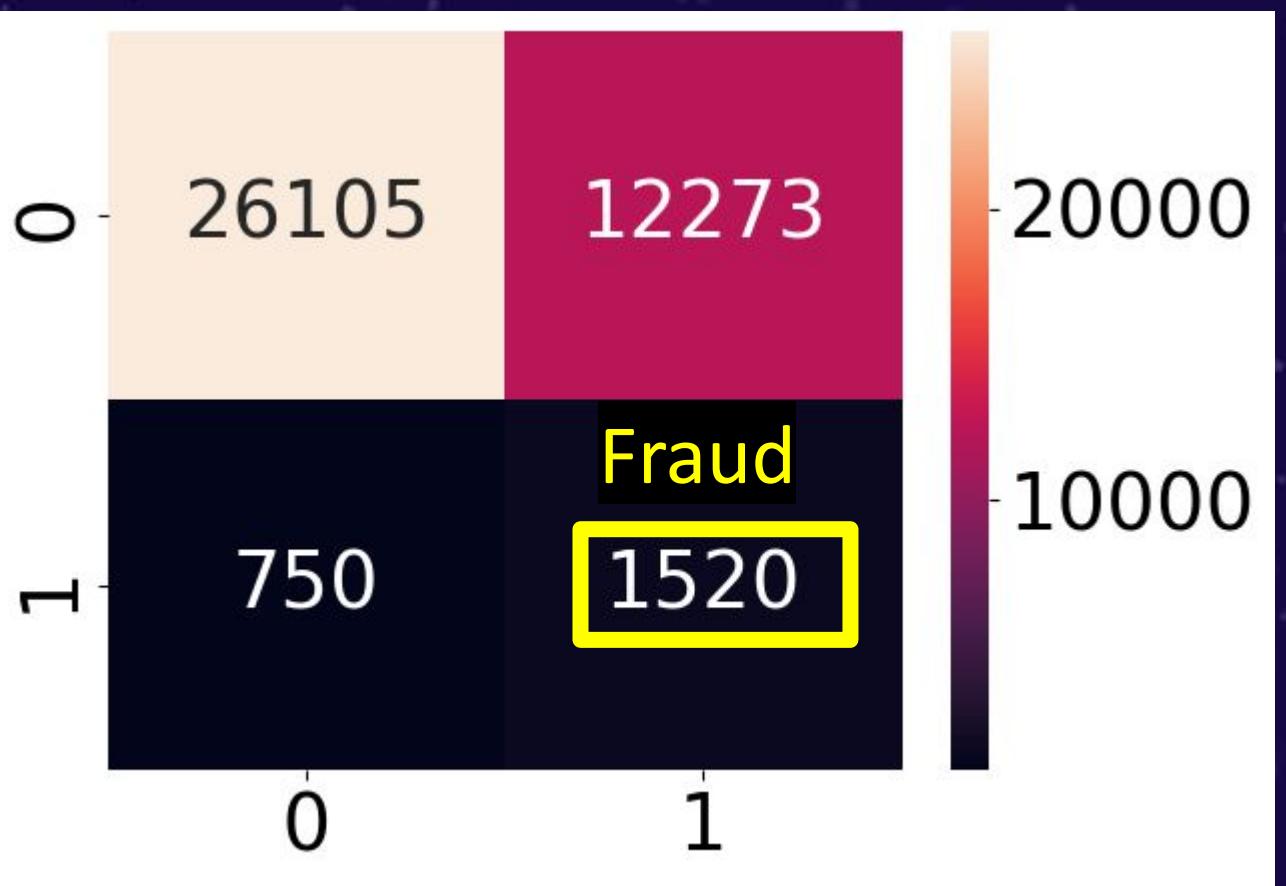
## Levels



# Model Performance

Our Model identifies a non-negligible proportion of Fraud!  
STEG will be able to tap into loss-prevention by revealing fraudulent customers

Potential recovered funds: 200.890 Euros per year.

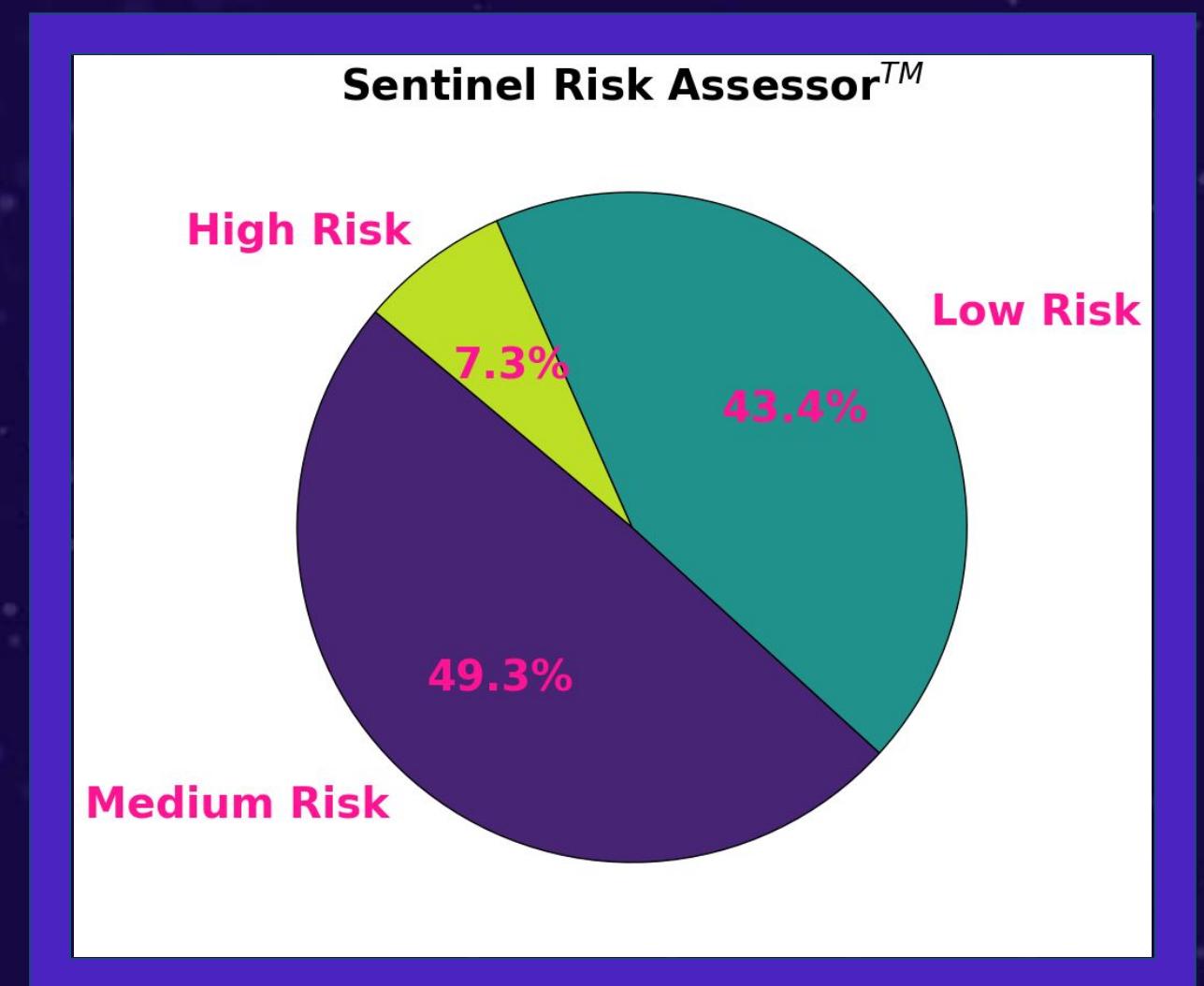


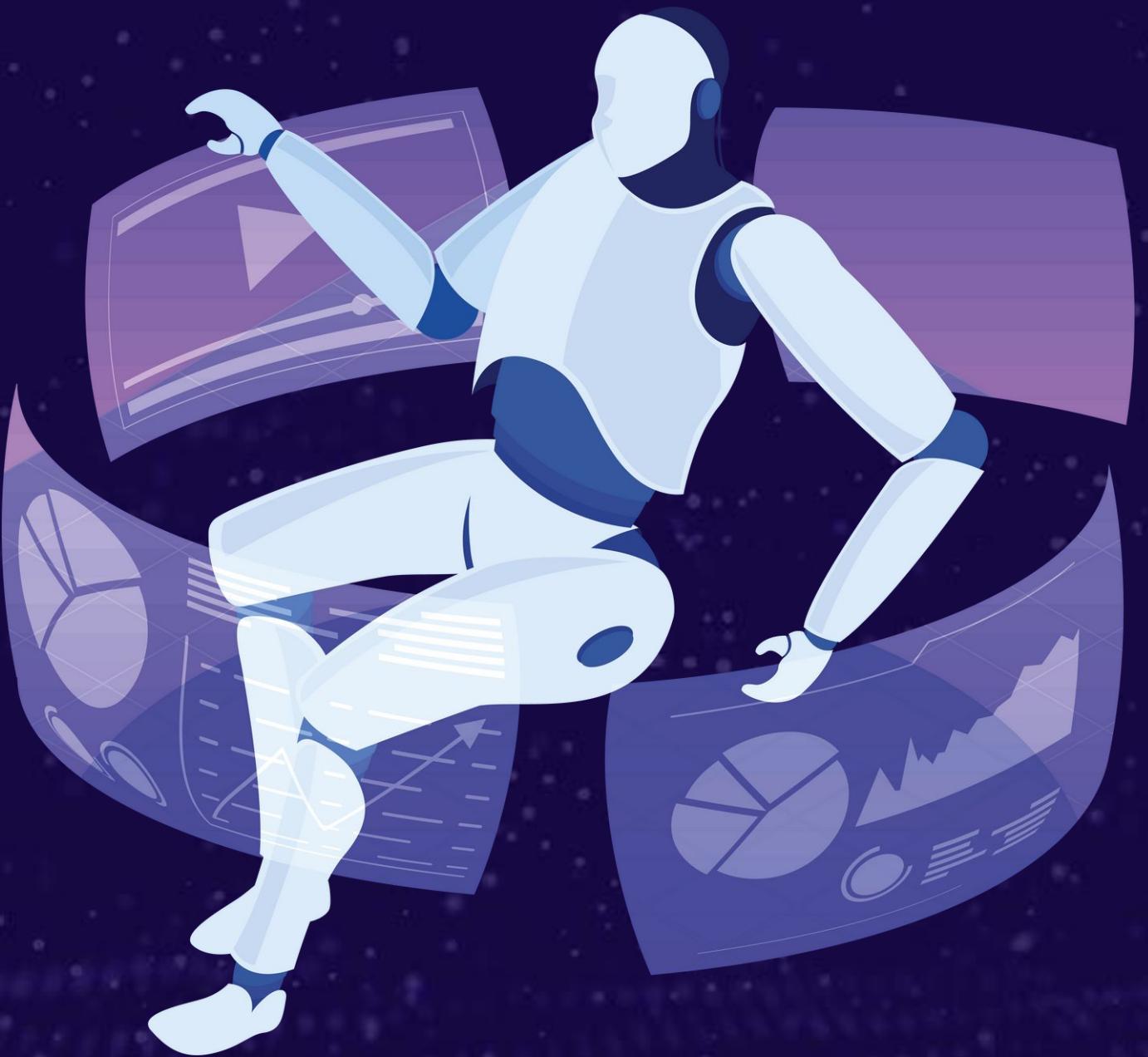
Threshold 1: 30%

flag and only use suspicious cases as 2nd layer input

Threshold 2: bins

- 0 - 40% : Low Risk
- 41 - 70% : Medium Risk
- 71 - 100% : High risk





Time series  
with tsfresh package

Feature Engineering

Ensemble methods

Optimized Data-Driven  
Fraud detection

# SunSafe Outlook

For STEG we have even more products and hold a variety of subscription plans ready!

We tailor and scale our models and services according to your needs.



SunSafe Repository

Data Source on Zindi

Presented by  
**Team Sunflower**



**Website**  
[www.SunSafe.com](http://www.SunSafe.com)

# Thank You!

Get in touch with us



[hi@SunSafe.com](mailto:hi@SunSafe.com)

## Feedback

Feature engineering: frequency of invoice\_date ( count) in aggregation

XGBoost does not need any feature encoding ( despite being tree based, it's a different package. scikit-learn needs encoding, xgboost not)

1. Business goal 2. highlight and stick to metric of choice for all models