Machine Learning pour la détection d'intrusion

Les bonnes pratiques

Anaël Beaugnon

anael.beaugnon@ssi.gouv.fr

ANSSI, ENS Paris, INRIA

Forum CERT IST 2017

L'intelligence artificielle et la sécurité

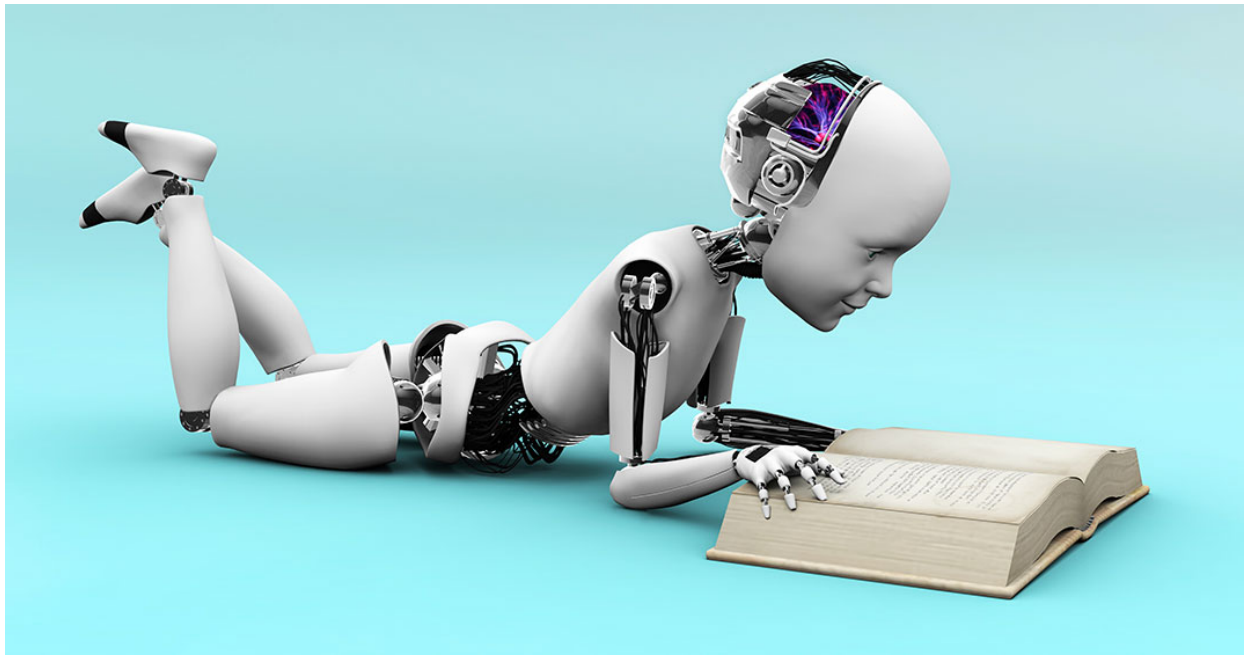
Sommaire

1 Contexte et problème

2 Machine Learning

3 Bien utiliser le Machine Learning !

4 Obtenir un bon jeu de données annotées



Intelligence Artificielle

Deep Learning

Data Science

Big Data

Machine Learning

Un succès dans de nombreux domaines

- Recommandation de produits sur Amazon
- Détection et reconnaissance de visages sur Facebook
- Intelligence artificielle pour le Go de Google (AlphaGo)

Machine Learning et détection d'intrusion

Les plaquettes marketing !

- ✓ Analyse comportementale
- ✓ Attaques inconnues
- ✓ 0-day

Pas si simple ...

- ✗ Trop de faux positifs
- ✗ Boîte noire incompréhensible

Comment adapter le Machine Learning à la détection d'intrusion ?

Sommaire

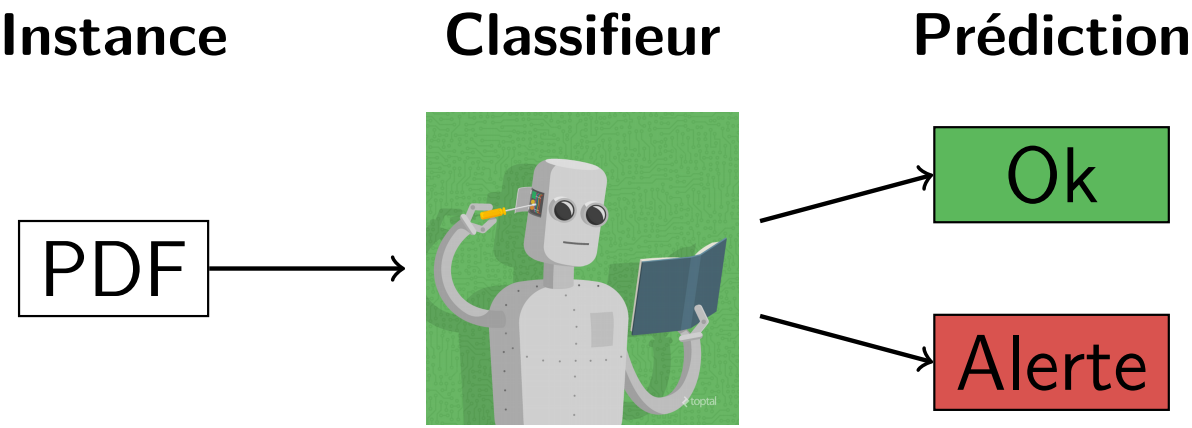
1 Contexte et problème

2 Machine Learning

3 Bien utiliser le Machine Learning !

4 Obtenir un bon jeu de données annotées

Machine Learning - Classifieur

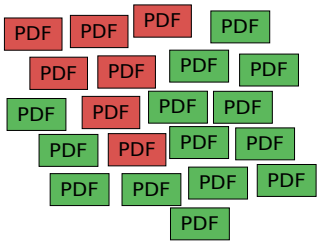


Deux étapes

1. Apprentissage du classifieur
2. Détection

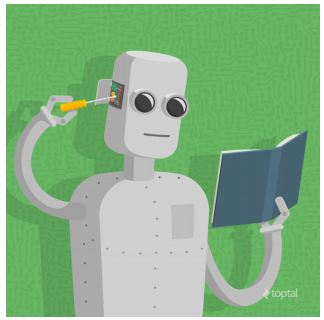
1- Apprentissage d'un classifieur

Données annotées

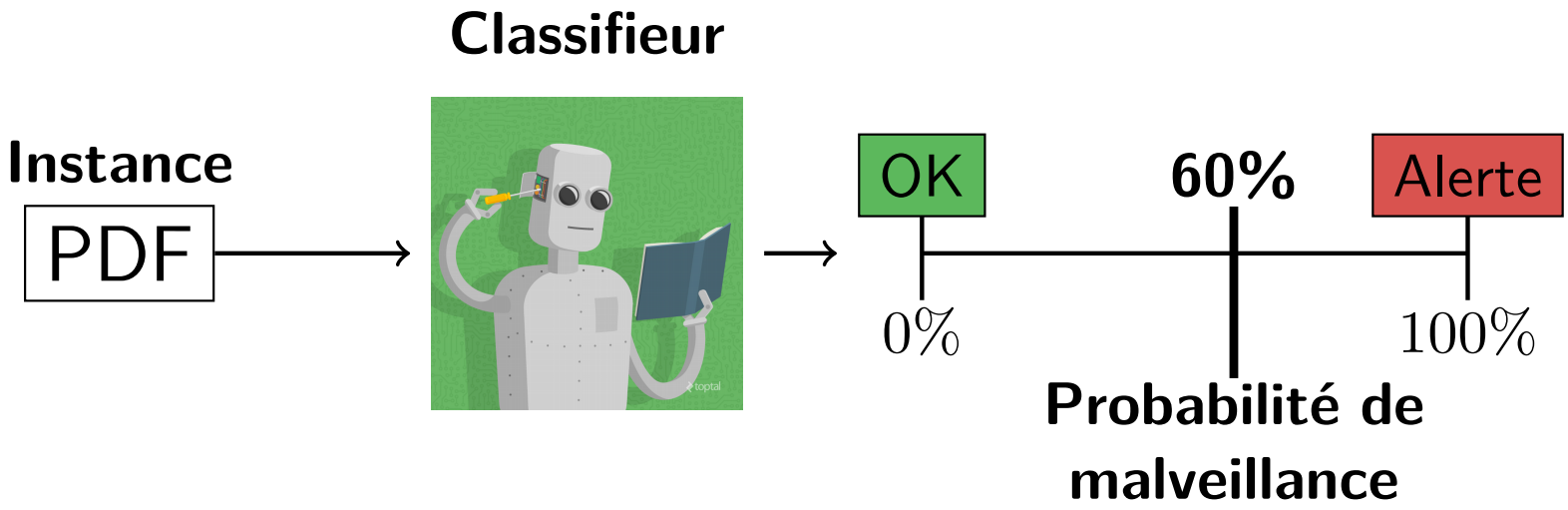


Algorithme d'apprentissage automatique

Classifieur



2- Détection grâce au classifieur

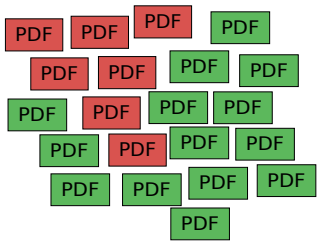


Probabilité de malveillance

- Prioritisation des alertes
- Compromis entre taux de détection et taux de faux positifs

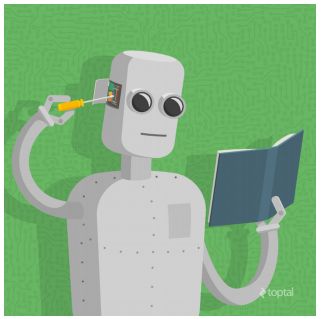
Vecteurs d'attributs numériques

Données brutes

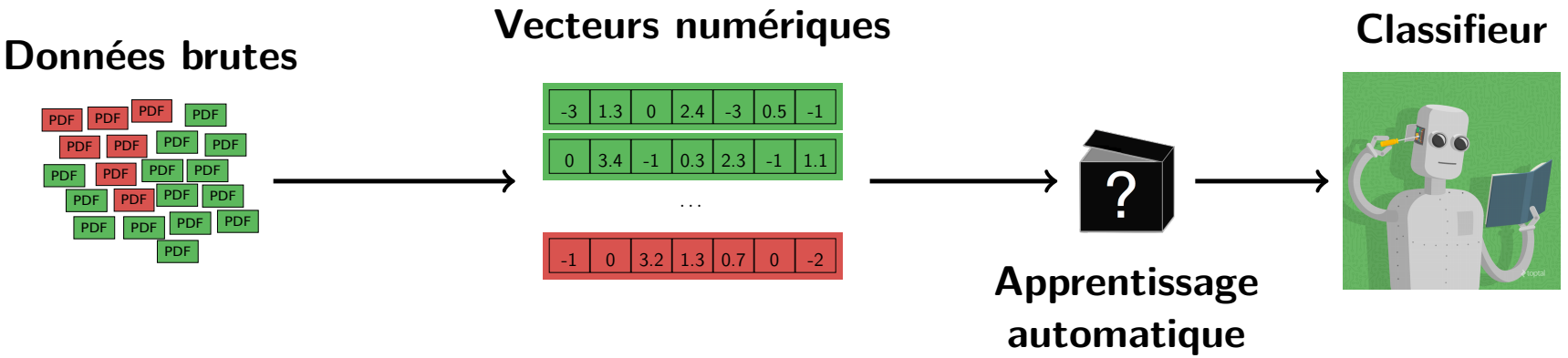


Apprentissage automatique

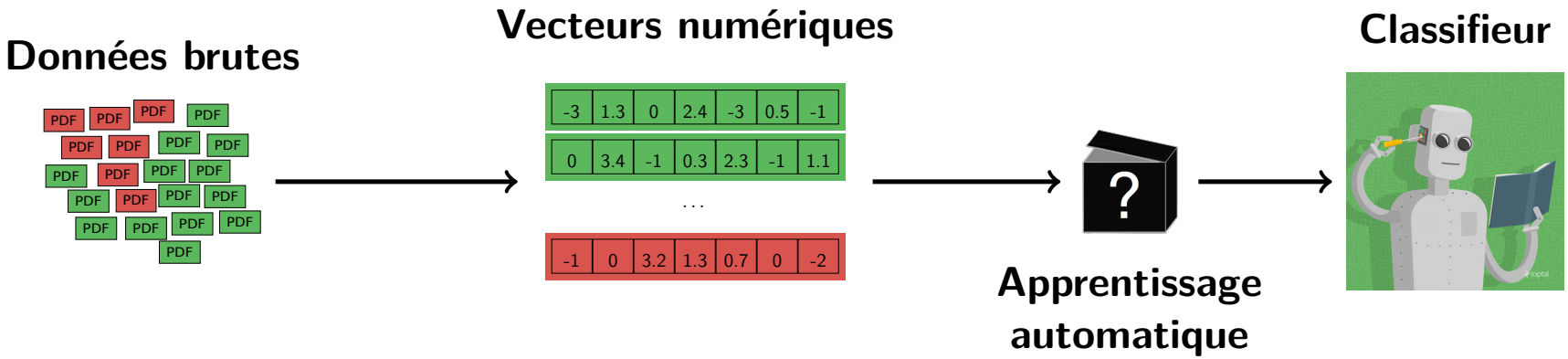
Classifieur



Vecteurs d'attributs numériques



Vecteurs d'attributs numériques



Extraction d'attributs

Donnée brute

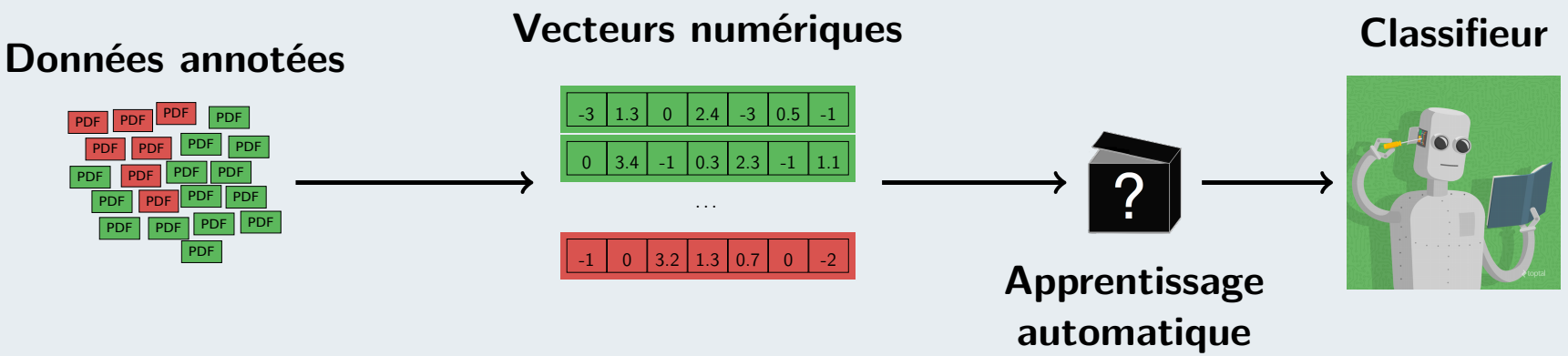
PDF

Vecteur numérique

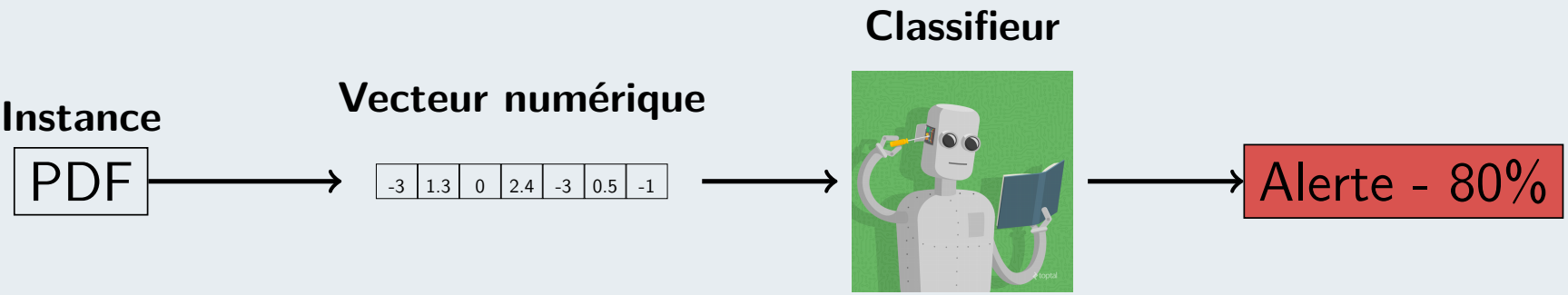
-3	1.3	0	2.4	-3	0.5	-1

Machine Learning

1- Apprentissage



2- Détection



Sommaire

1 Contexte et problème

2 Machine Learning

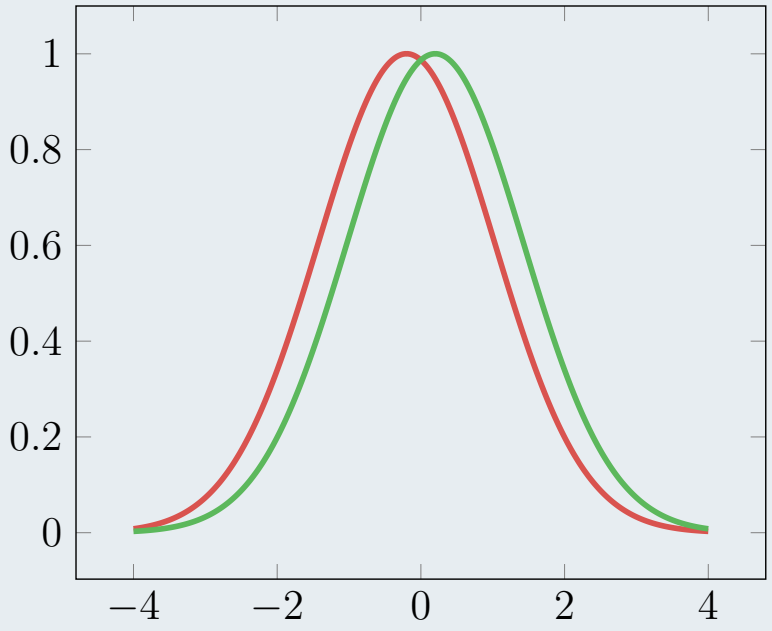
3 Bien utiliser le Machine Learning !

4 Obtenir un bon jeu de données annotées

1- Attributs discriminants

- Spécifiques à chaque problème de détection
- Connaissances expert

Non discriminant



Discriminant

2- Choix du modèle

Contraintes opérationnelles

- Prédiction rapide
- Mise à jour périodique du modèle
- Transparence - interprétation du modèle

2- Choix du modèle

Contraintes opérationnelles

- Prédiction rapide
- Mise à jour périodique du modèle
- Transparence - interprétation du modèle

✗ Réseaux de neurones

2- Choix du modèle

Contraintes opérationnelles

- Prédiction rapide
- Mise à jour périodique du modèle
- Transparence - interprétation du modèle

✗ Réseaux de neurones

✗ k plus proches voisins

2- Choix du modèle

Contraintes opérationnelles

- Prédiction rapide
- Mise à jour périodique du modèle
- Transparence - interprétation du modèle

✗ Réseaux de neurones

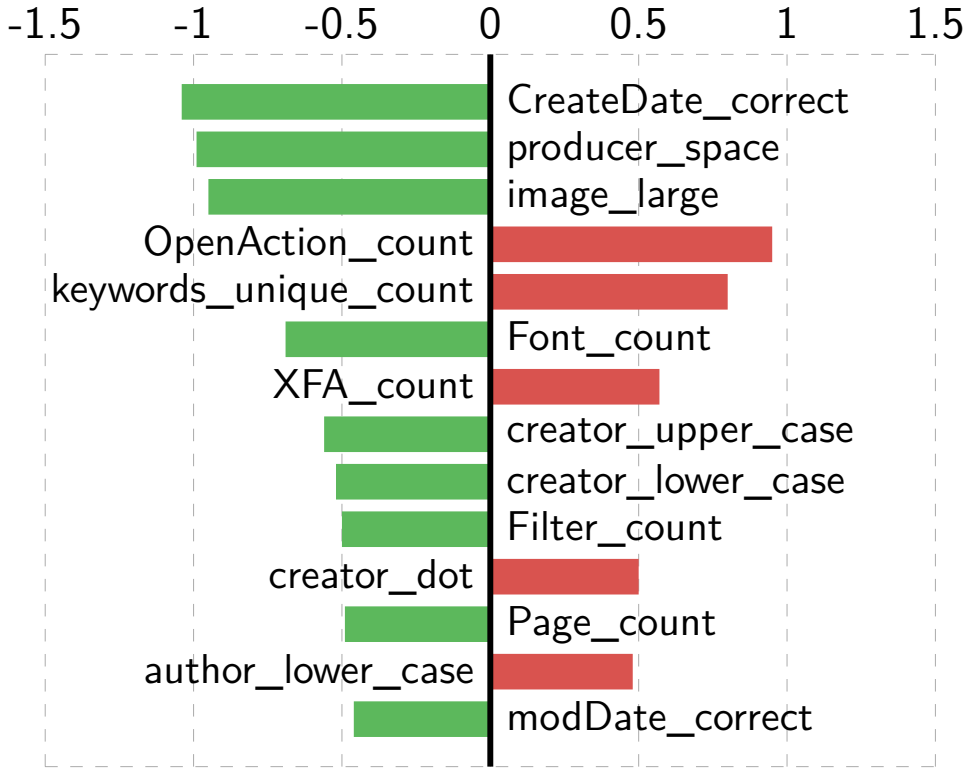
✗ k plus proches voisins

✓ Modèles linéaires (ex: régression logistique, SVM)

✓ Modèles d'arbre (ex: arbre de décision, forêt aléatoire)

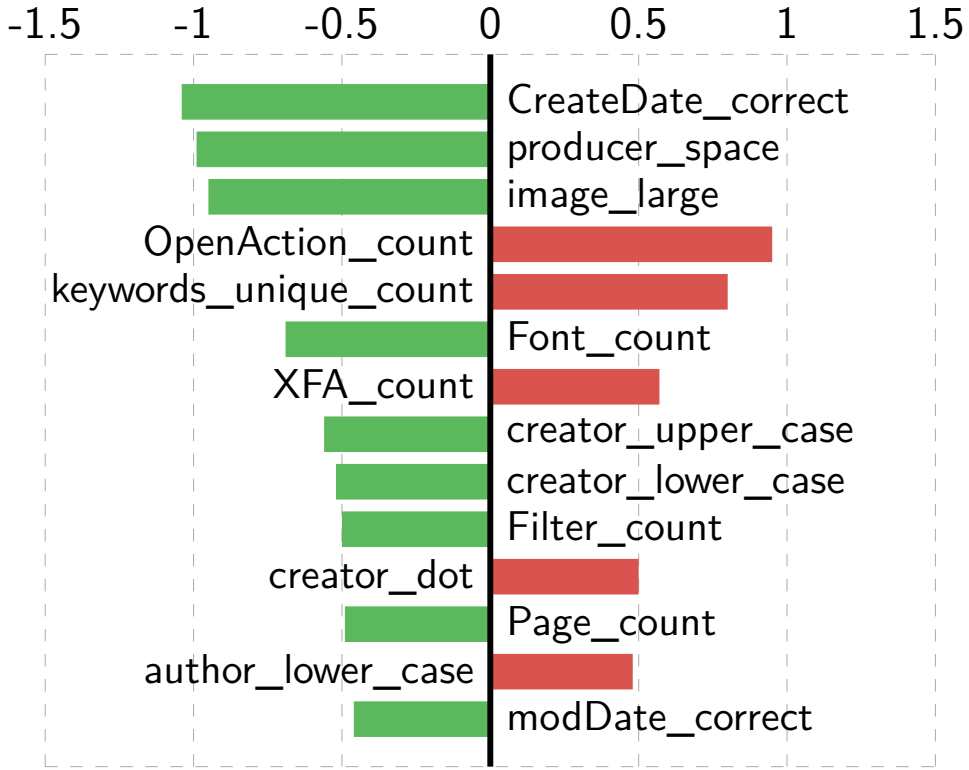
2- Choix du modèle

Les modèles linéaires sont interprétables.



2- Choix du modèle

Les modèles linéaires sont interprétables.



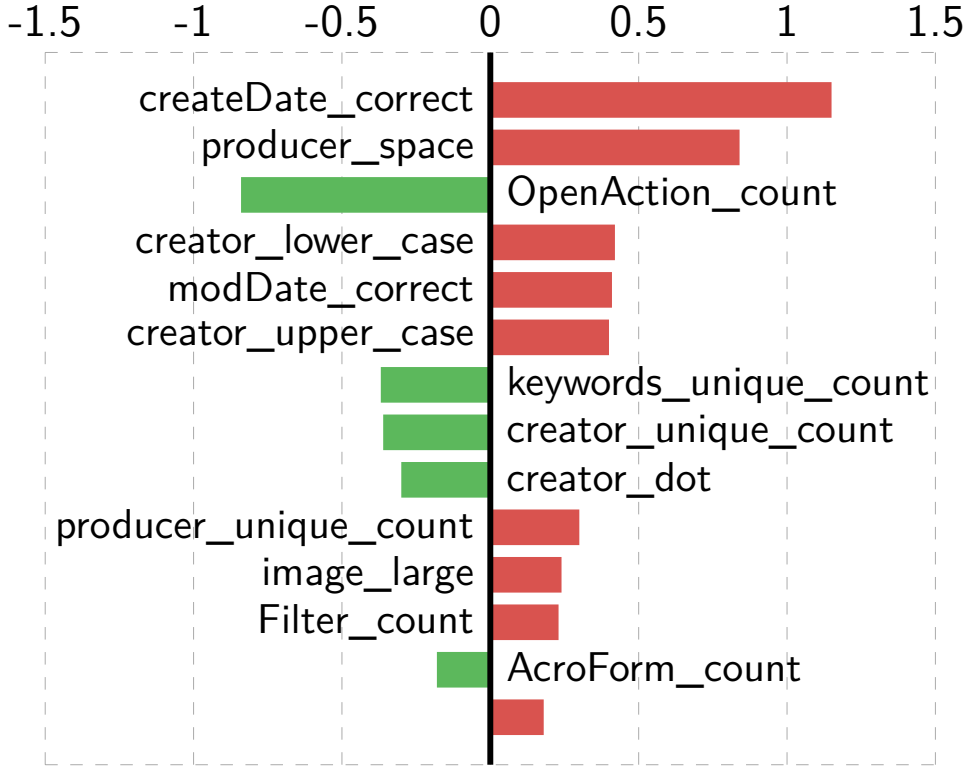
Méthode de scoring

Coefficients optimaux appris automatiquement à partir des données annotées

2- Choix du modèle

Les prédictions sont aussi interprétables !

Pourquoi une alerte a été générée ?



3- Valider le modèle

Avant la mise en production !

Jeu de données de validation

- Données annotées
- Validation sur des données non utilisées pour l'apprentissage

Méthode de validation

Apprentissage	Validation
90% données	**10% données**

Bien utiliser le Machine Learning !

Bonnes pratiques

1. Attributs discriminants
2. Modèle répondant aux contraintes opérationnelles
3. Validation du modèle

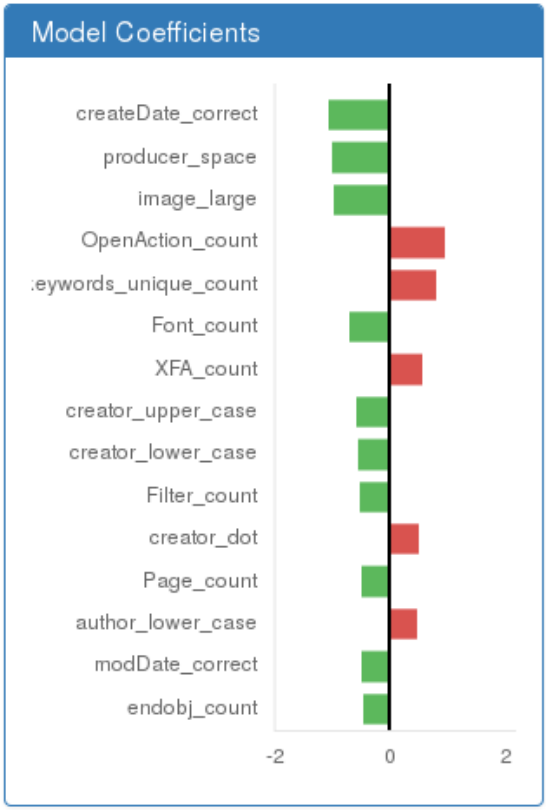
SecuML

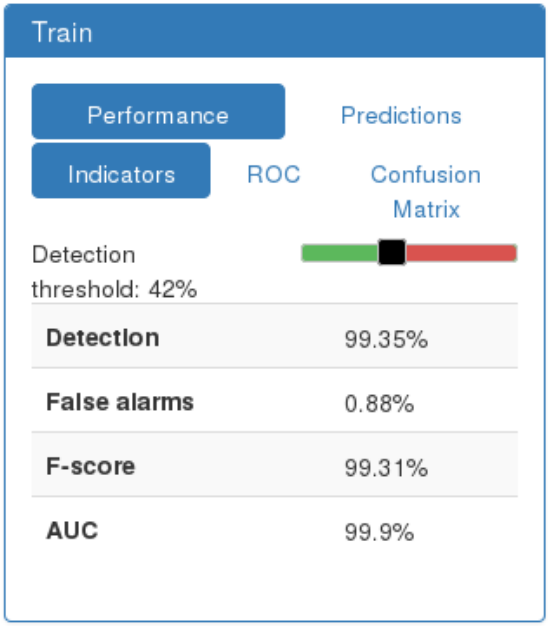
- Interface de diagnostic d'un classifieur
- Mise en place d'un modèle avant sa mise en production
- https://github.com/ANSSI-FR/SecuML

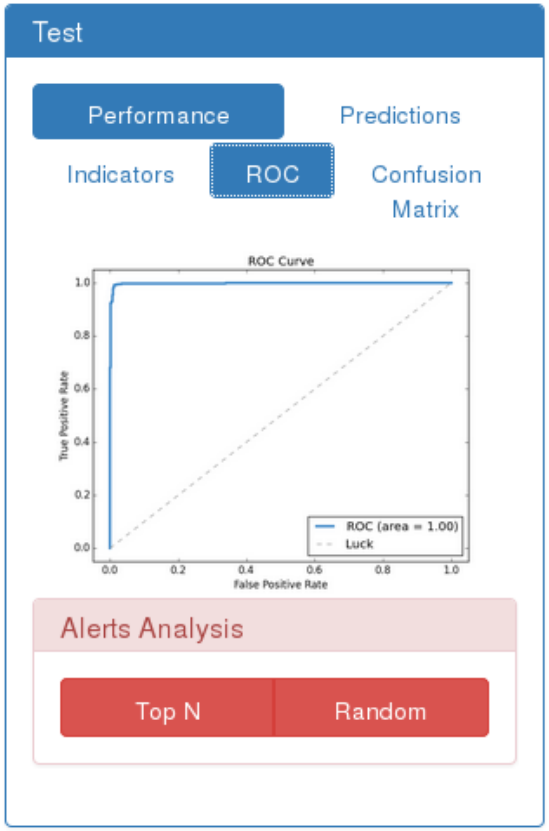
SSTIC 2017 Bonneton et al., Le Machine Learning confronté aux contraintes opérationnelles des systèmes de détection.

SecuML

Interface de diagnostic d'un classifieur







https://github.com/ANSSI-FR/SecuML

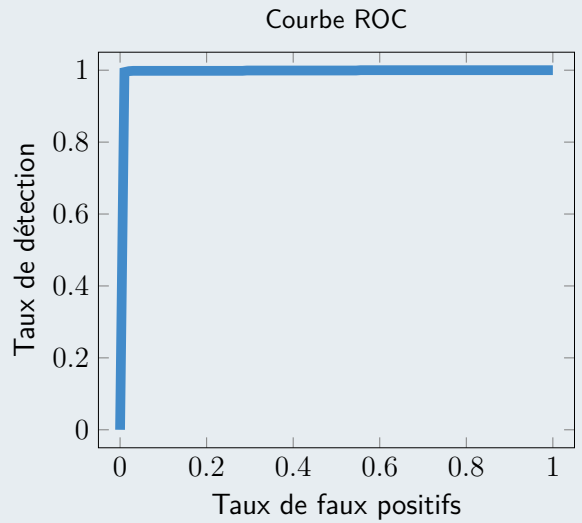
Détection de fichiers PDF malveillants

1. **Attributs:** de nombreux articles de recherche
2. **Modèle:** régression logistique
3. **Validation:** Contagio et WebPdf

Détection de fichiers PDF malveillants

1. **Attributs:** de nombreux articles de recherche
2. **Modèle:** régression logistique
3. **Validation:** Contagio et WebPdf

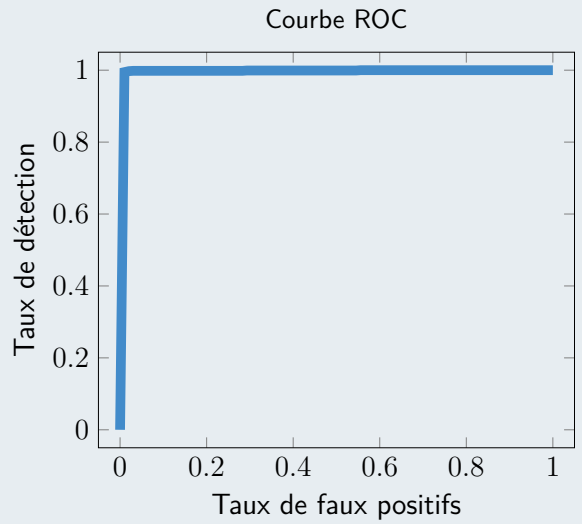
90% Contagio/ 10% Contagio



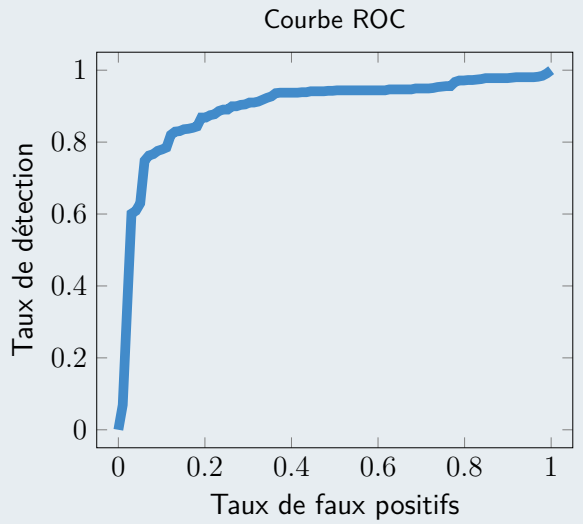
Détection de fichiers PDF malveillants

1. **Attributs:** de nombreux articles de recherche
2. **Modèle:** régression logistique
3. **Validation:** Contagio et WebPdf

90% Contagio/ 10% Contagio



100% Contagio/ 100% WebPdf



Bien utiliser le Machine Learning !

0. Un bon jeu de données annotées
1. Attributs discriminants
2. Modèle répondant aux contraintes opérationnelles
3. Validation du modèle

Bien utiliser le Machine Learning !

0. Un bon jeu de données annotées
1. Attributs discriminants
2. Modèle répondant aux contraintes opérationnelles
3. Validation du modèle

Comment obtenir un bon jeu de données annotées ?

Sommaire

1 Contexte et problème

2 Machine Learning

3 Bien utiliser le Machine Learning !

4 Obtenir un bon jeu de données annotées

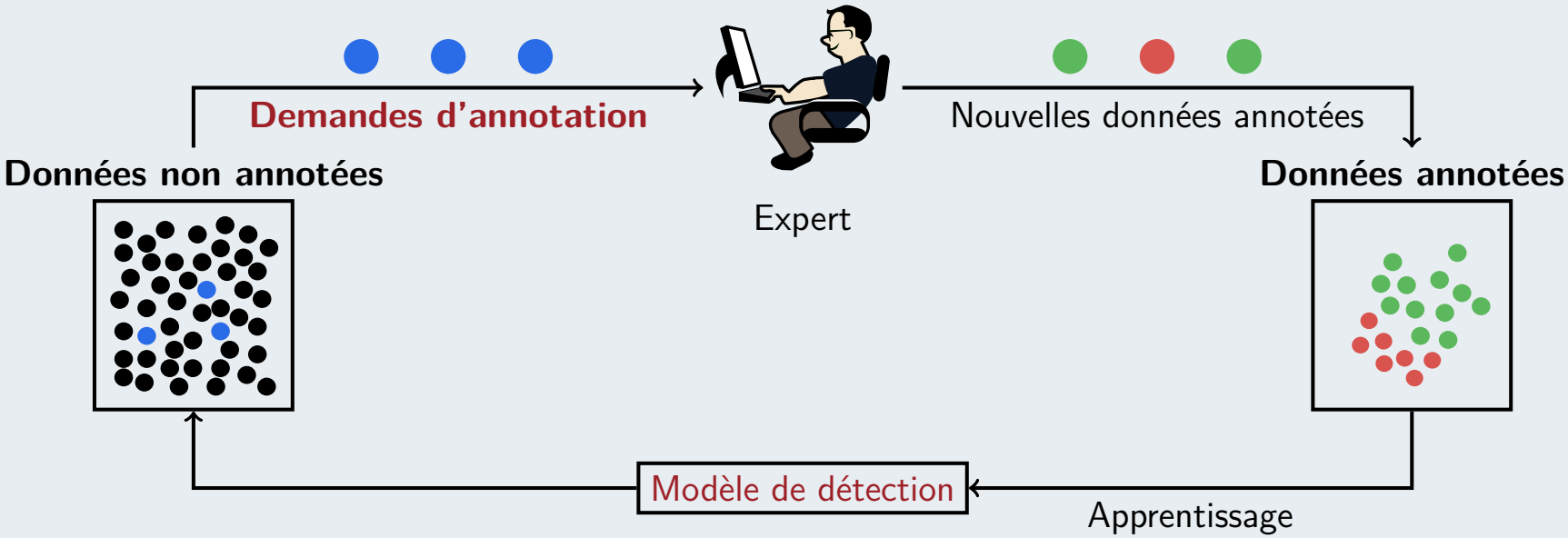
Manque de données d'apprentissage

- ✗ Jeux de données publics $\neq$ production
- ✗ Crowd-sourcing

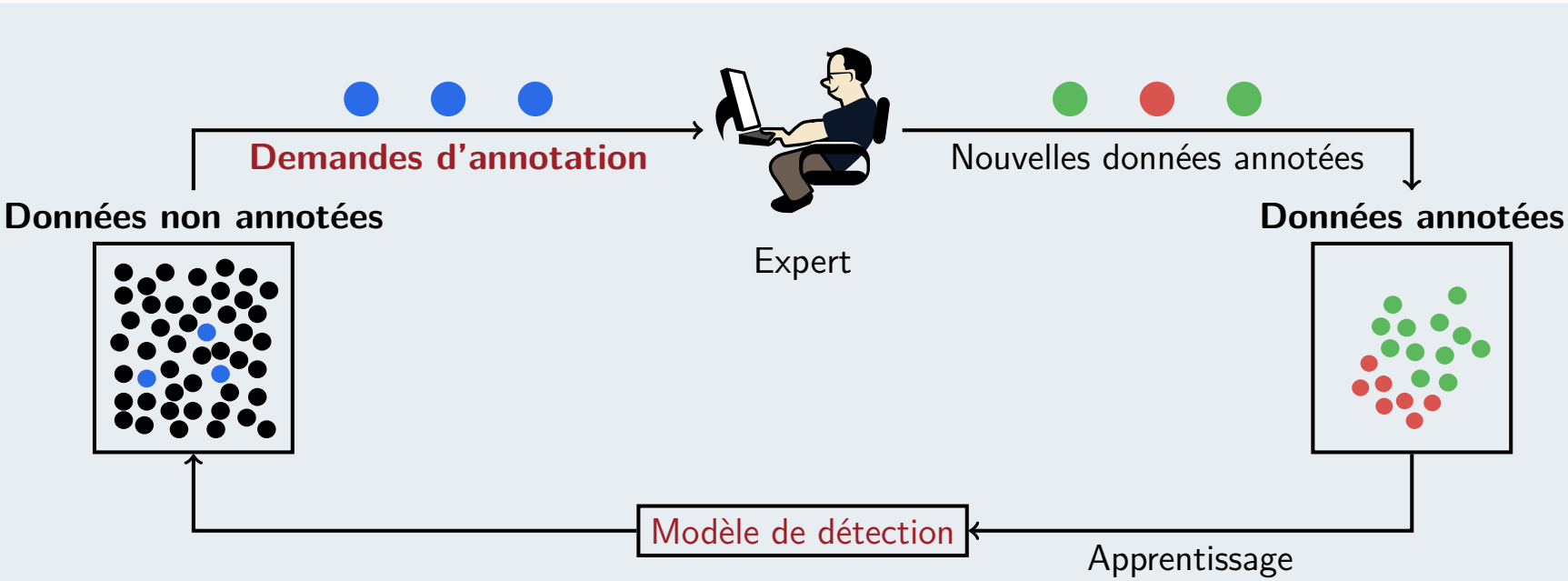
Manque de données d'apprentissage

- ✗ Jeux de données publics $\neq$ production
- ✗ Crowd-sourcing

Solution : Annotation in-situ



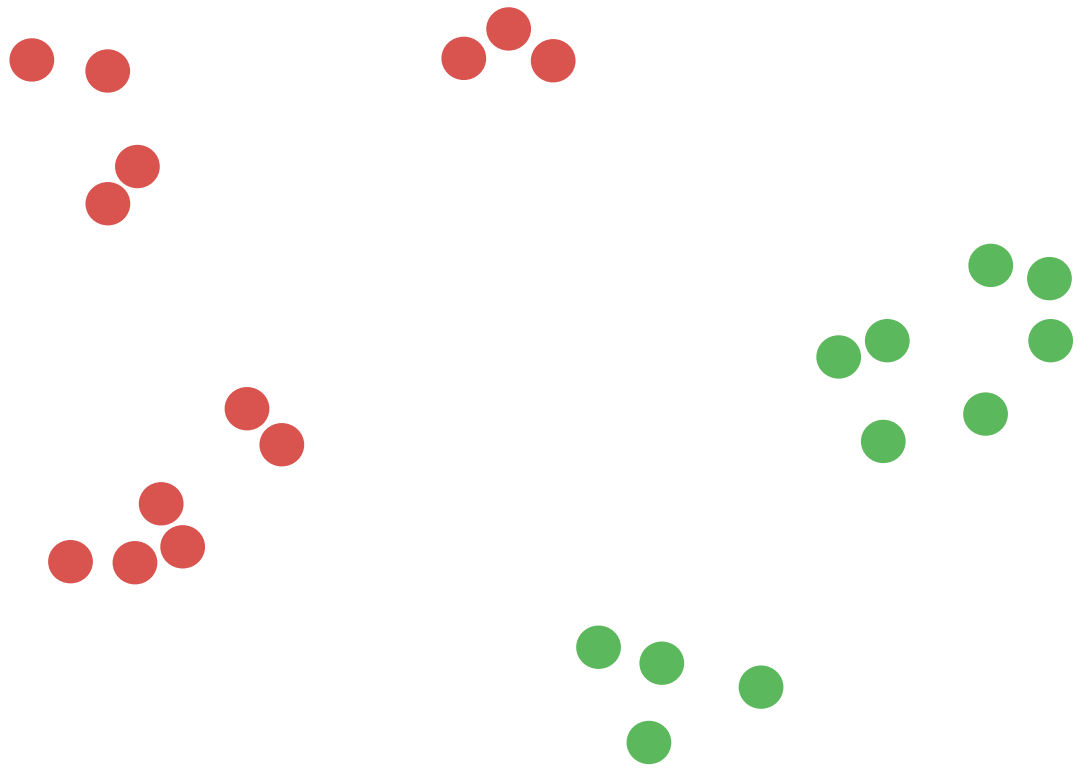
Comment sélectionner les demandes d'annotations ?



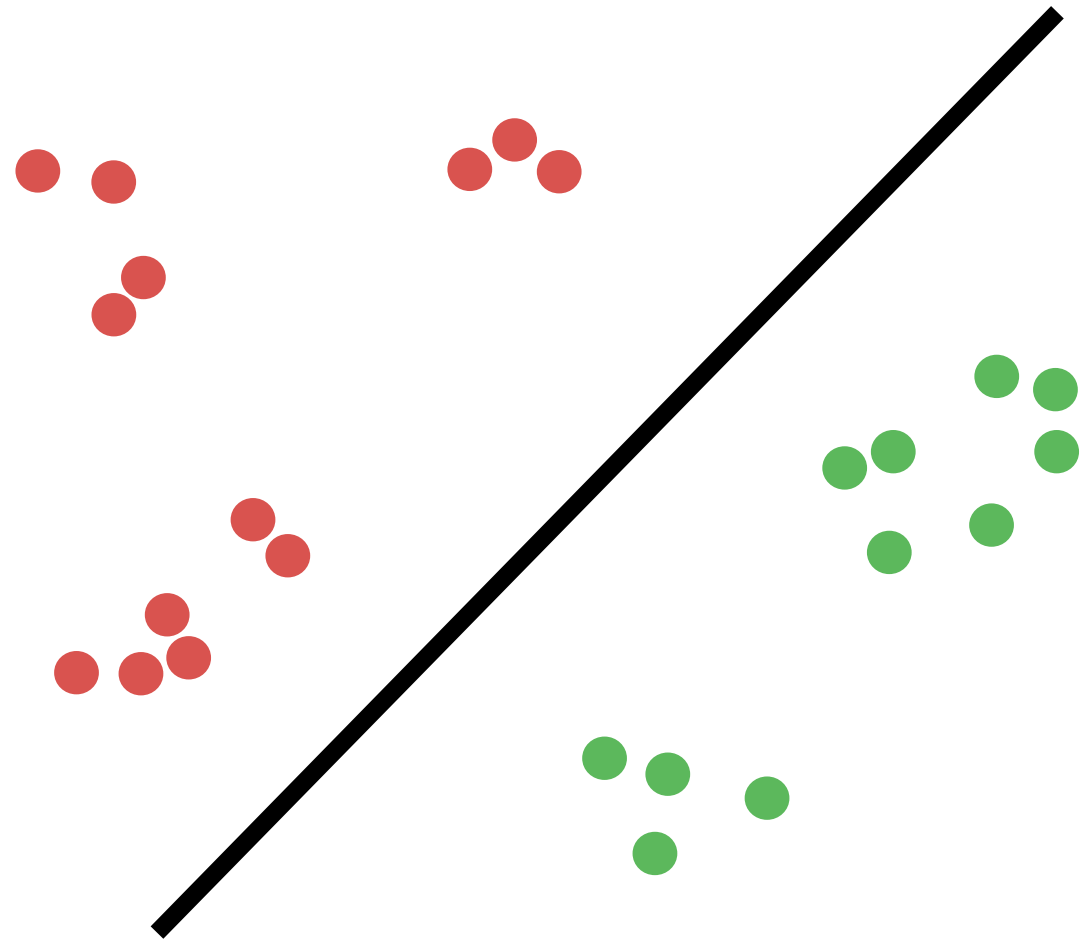
Méthode de sélection

- ✗ Sélection aléatoire
- ✓ Active learning

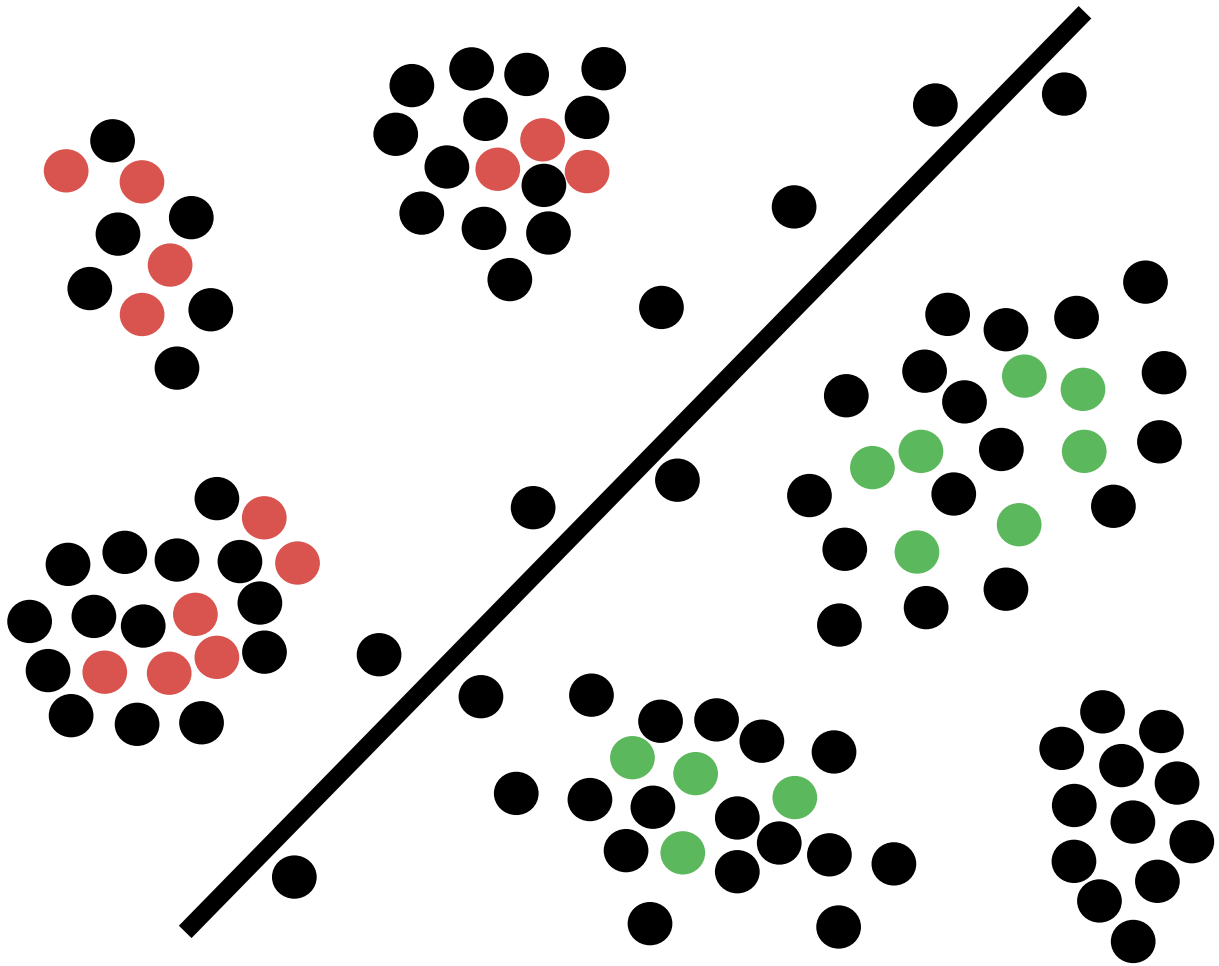
Principe de l'active learning (uncertainty sampling)



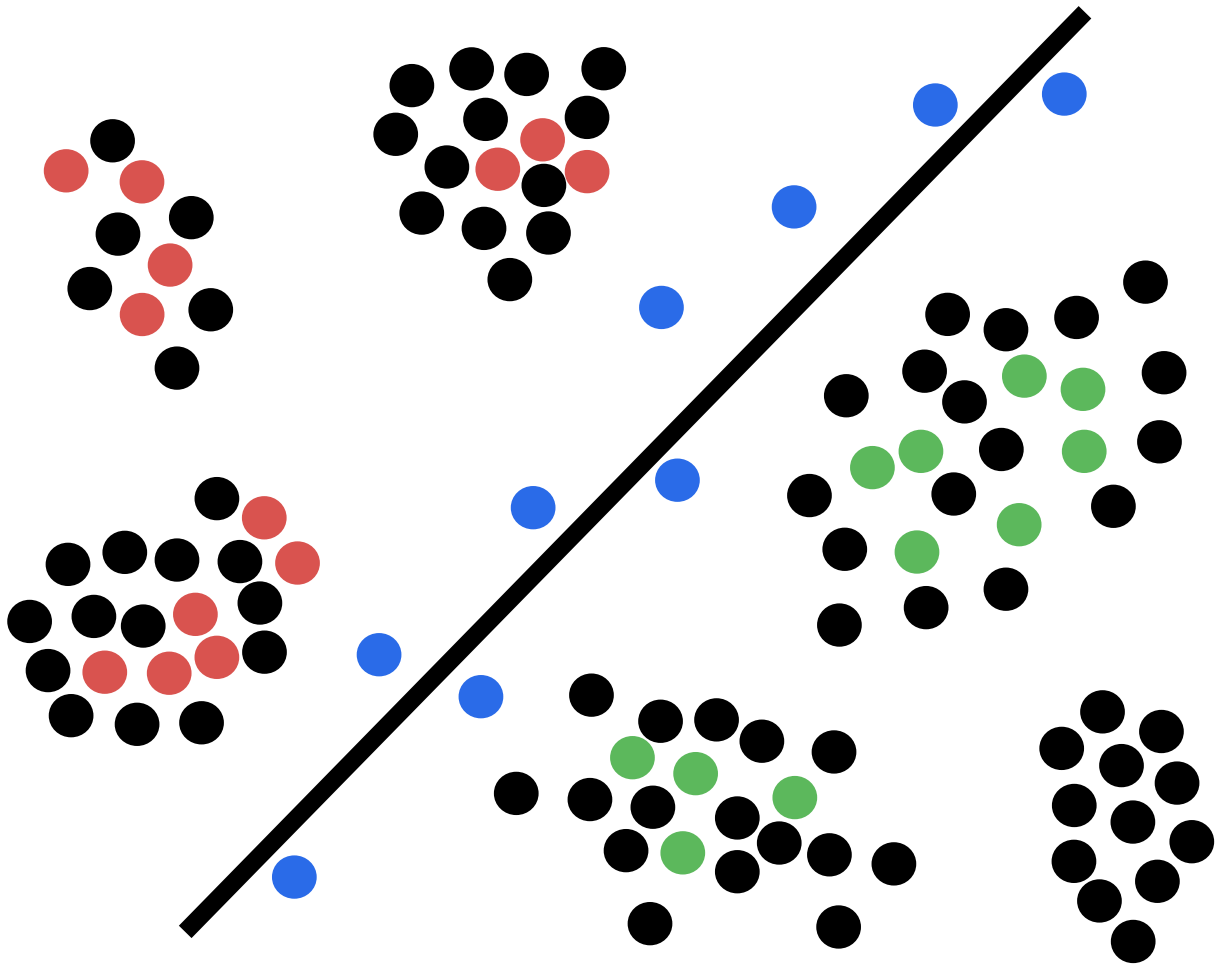
Principe de l'active learning (uncertainty sampling)



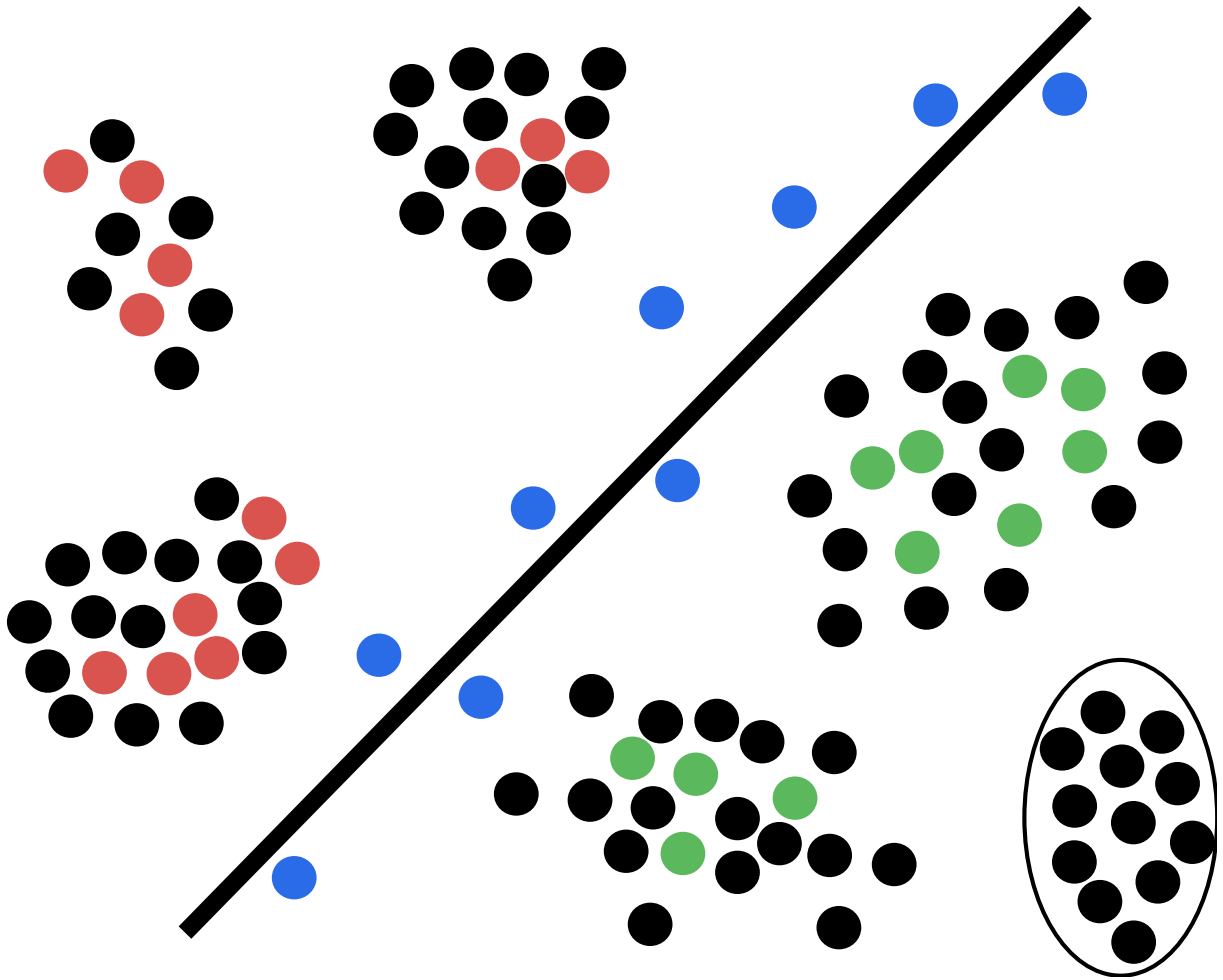
Principe de l'active learning (uncertainty sampling)



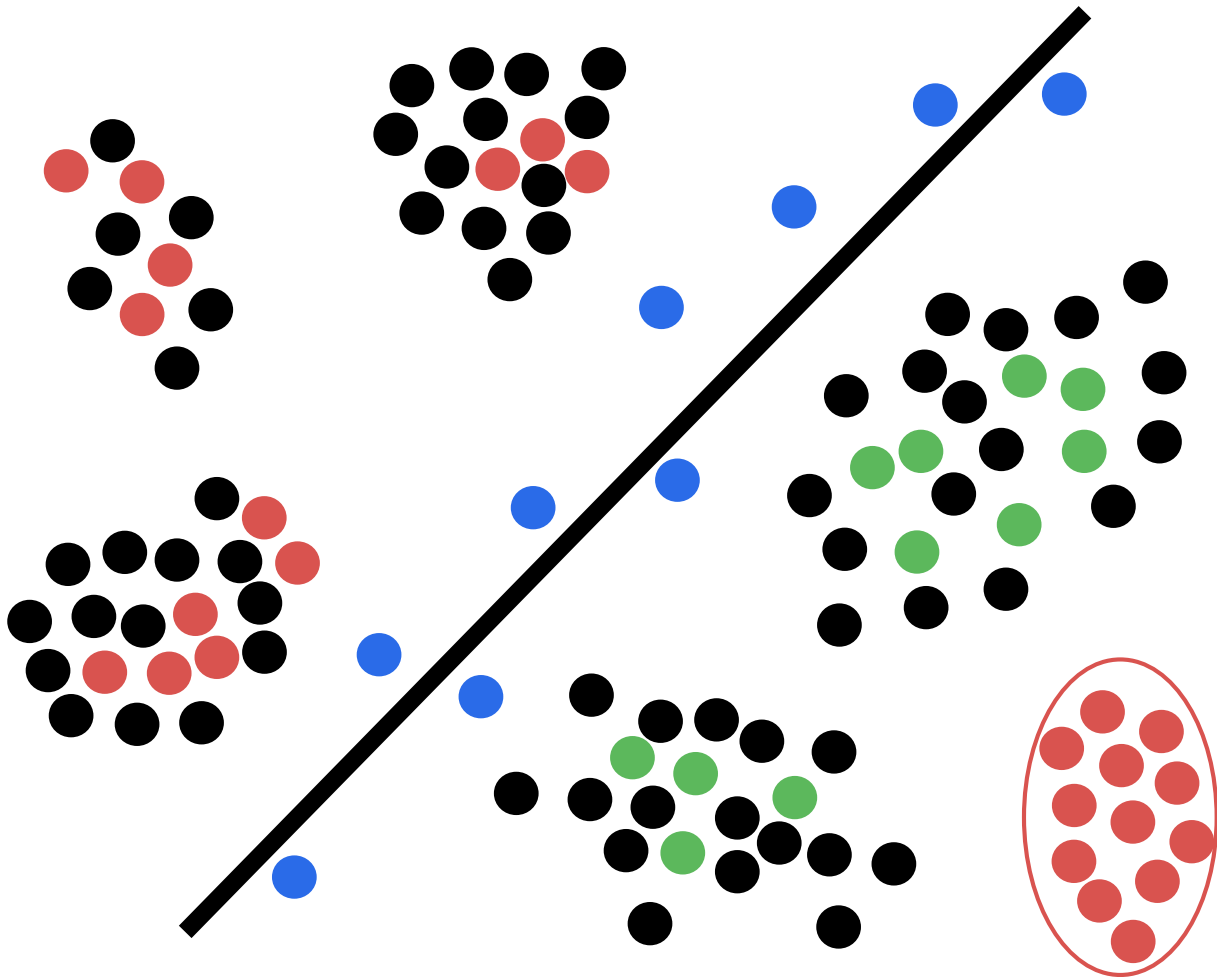
Principe de l'active learning (uncertainty sampling)



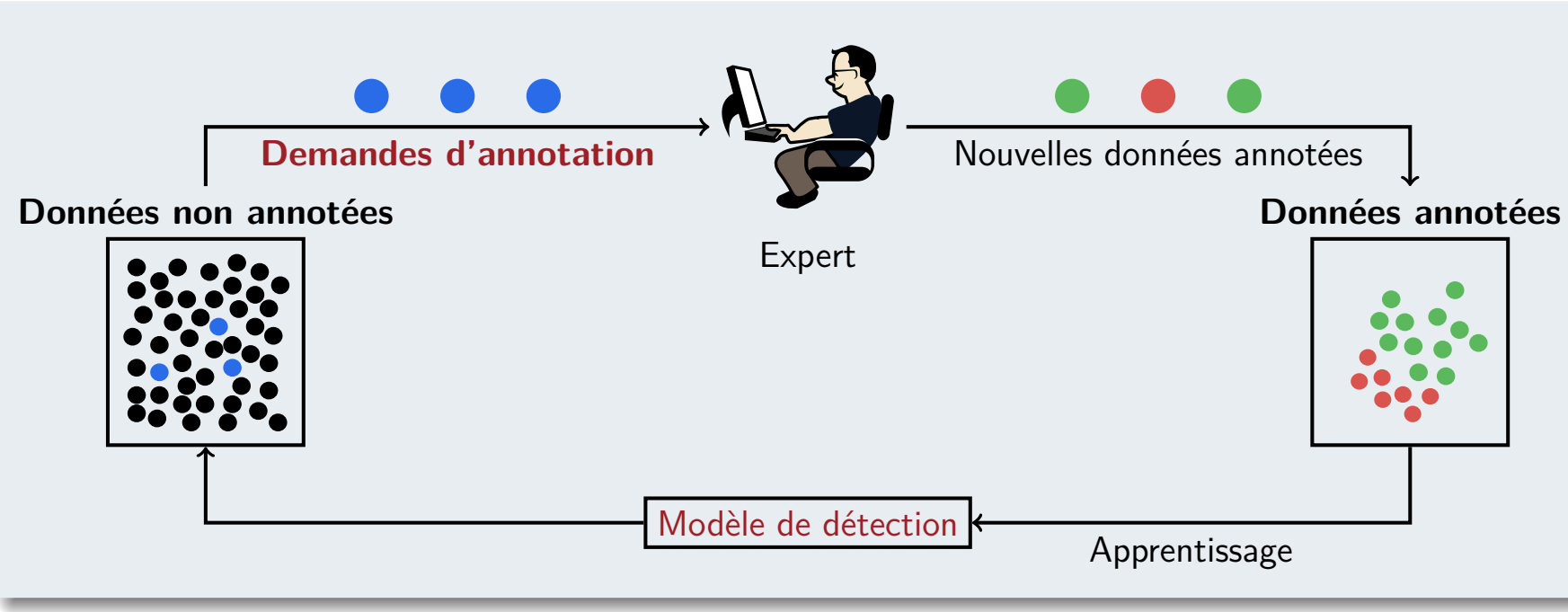
Principe de l'active learning (uncertainty sampling)



Principe de l'active learning (uncertainty sampling)



Défis liés à l'active learning



Défis

1. Détecter toutes les familles
2. Réduire le temps d'attente
3. Interface utilisateur adaptée

ILAB

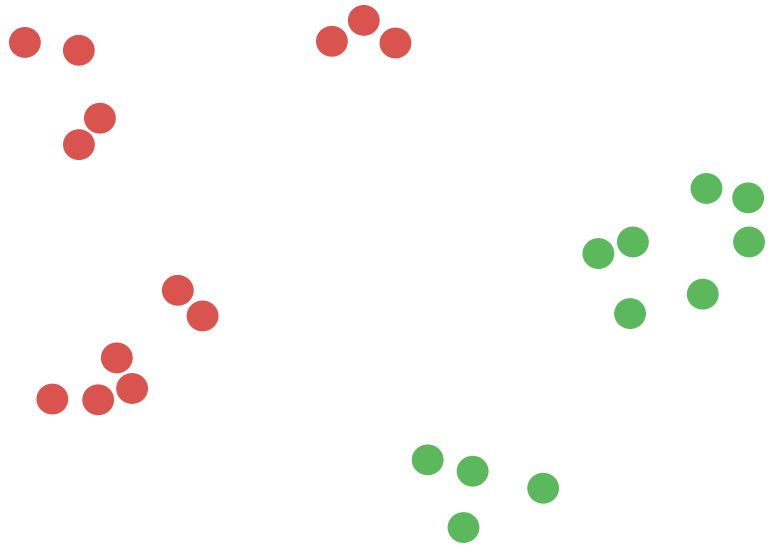
Un système d'annotation adapté aux besoins des experts en sécurité

Répond aux défis

1. Détecter toutes les familles
2. Réduire le temps d'attente
3. Interface utilisateur adaptée

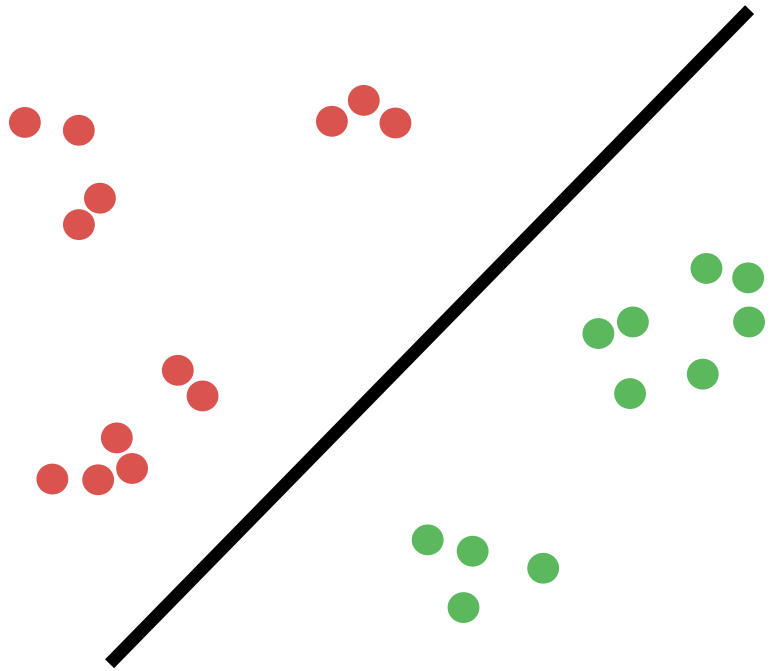
1- Détecter toutes les familles

Annotation : famille



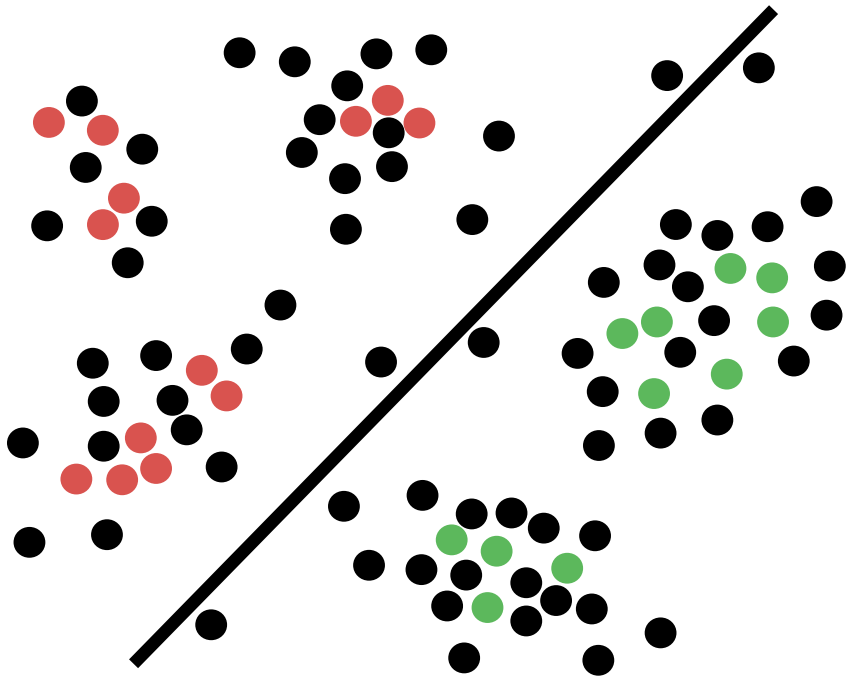
1- Détecter toutes les familles

Annotation : famille



1- Détecter toutes les familles

Annotation : famille

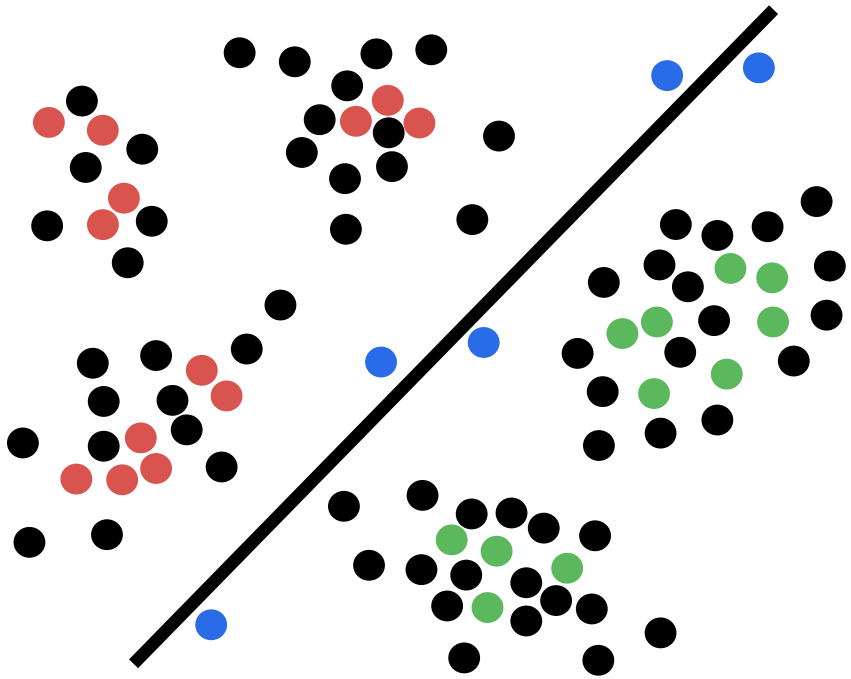


1- Détecter toutes les familles

Annotation : famille

Demande d'annotation

- Frontière de décision

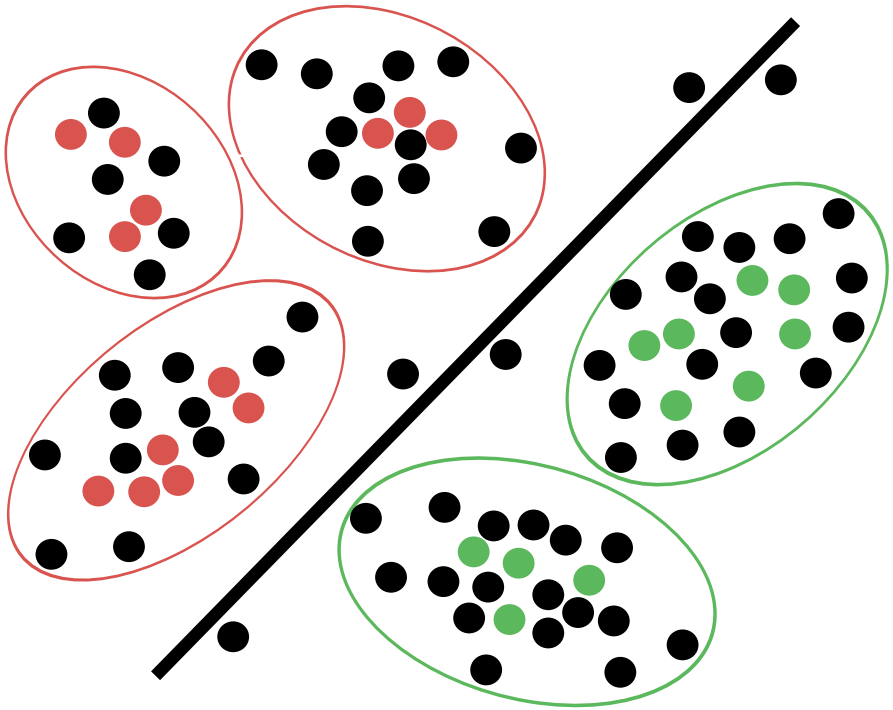


1- Détecter toutes les familles

Annotation : famille

Demande d'annotation

- Frontière de décision



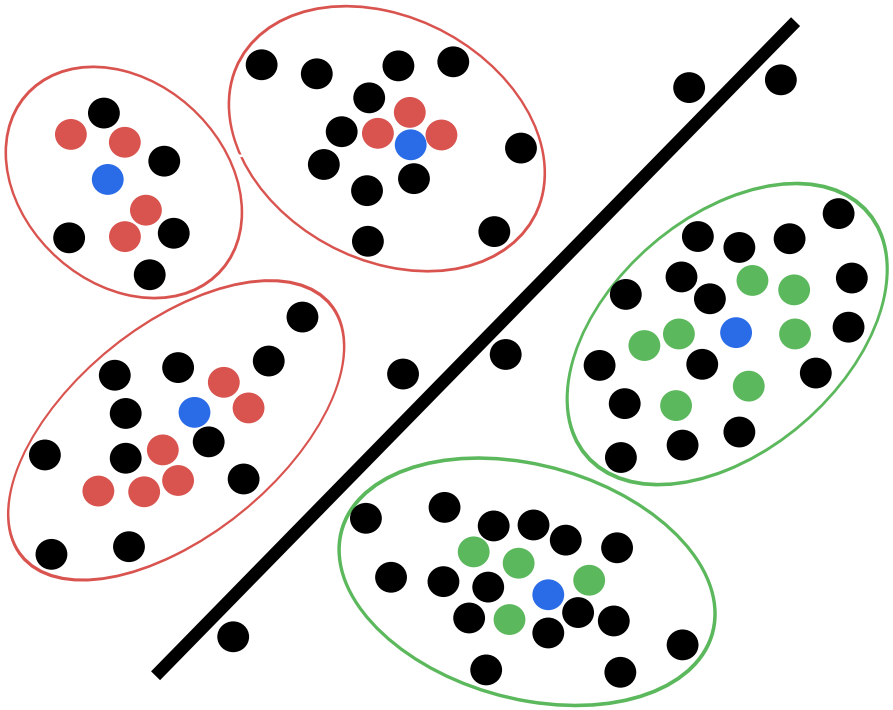
Clusters = Familles définies par l'utilisateur

1- Détecter toutes les familles

Annotation : famille

Demande d'annotation

- Frontière de décision
- Centre des clusters



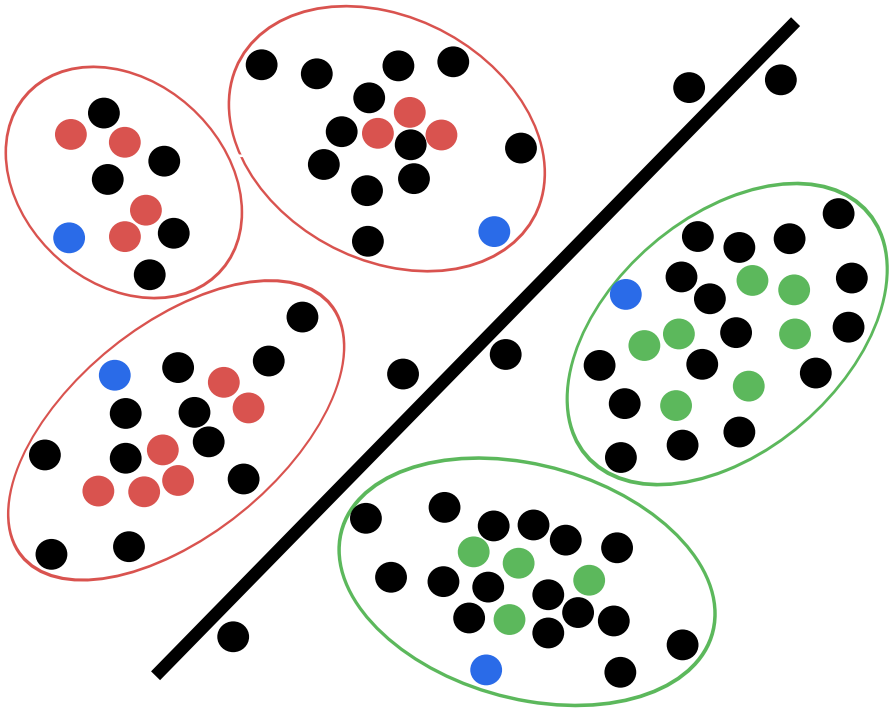
Clusters = Familles définies par l'utilisateur

1- Détecter toutes les familles

Annotation : famille

Demande d'annotation

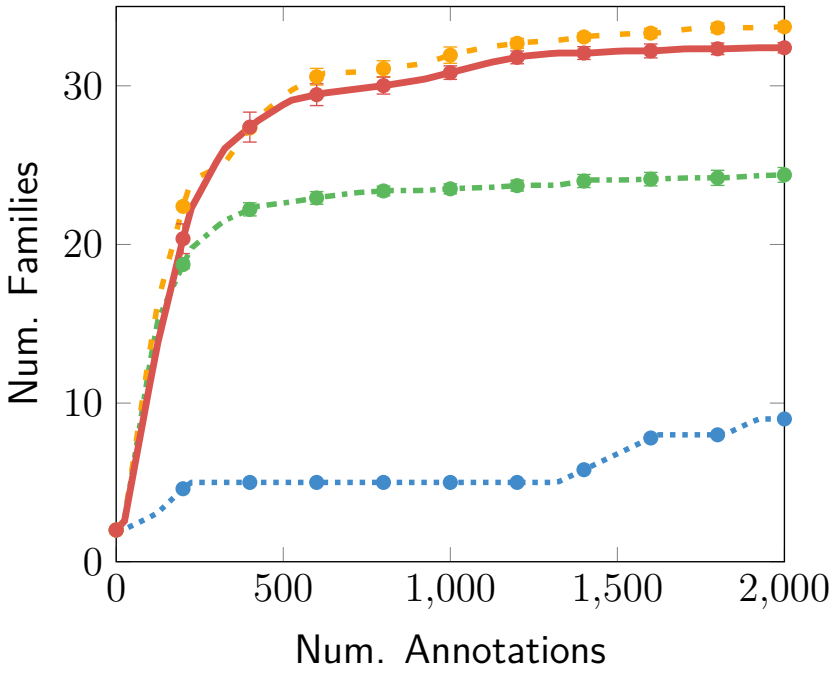
- Frontière de décision
- Centre des clusters
- Bord des clusters

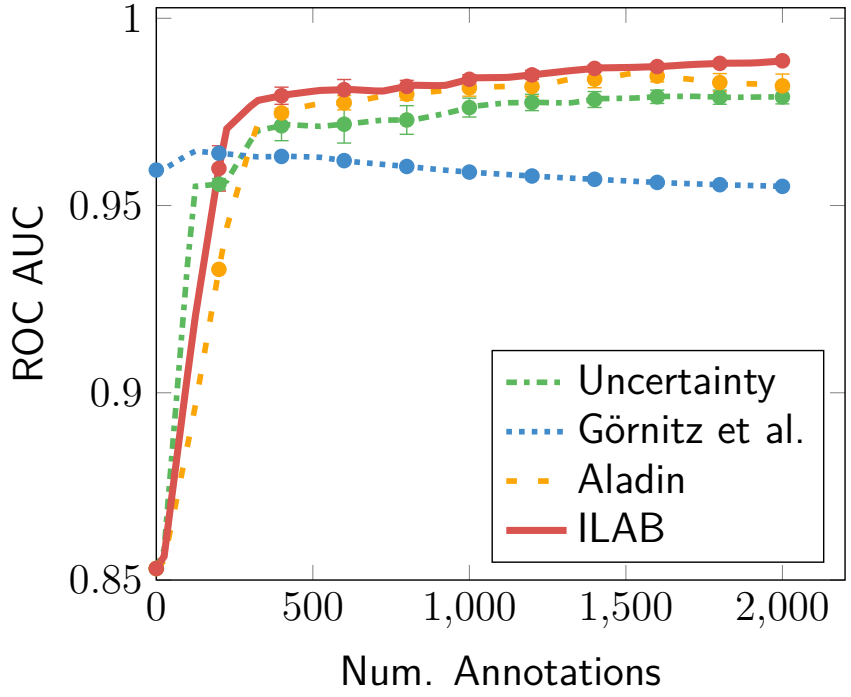


Clusters = Familles définies par l'utilisateur

1- Détecter toutes les familles

ILAB et Aladin détectent bien les différentes familles.



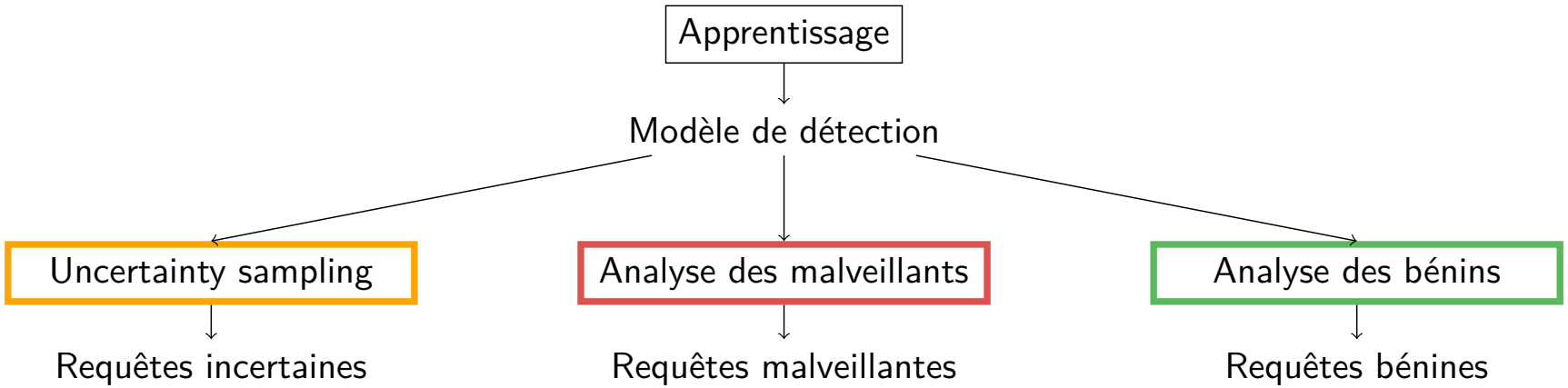


Uncertainty — Almgren et al., Using Active Learning in Intrusion Detection, CSFW 2004
Görnitz et al. — Görnitz et al., Toward Supervised Anomaly Detection, JAIR 2013
Aladin — Stokes et al., Aladin: Active Learning of Anomalies to Detect Intrusions, 2008
ILAB — Beaugnon et al., ILAB: An Interactive Labelling Strategy for Intrusion Detection, RAID 2017

2- Réduire le temps d'attente

Diviser pour régner

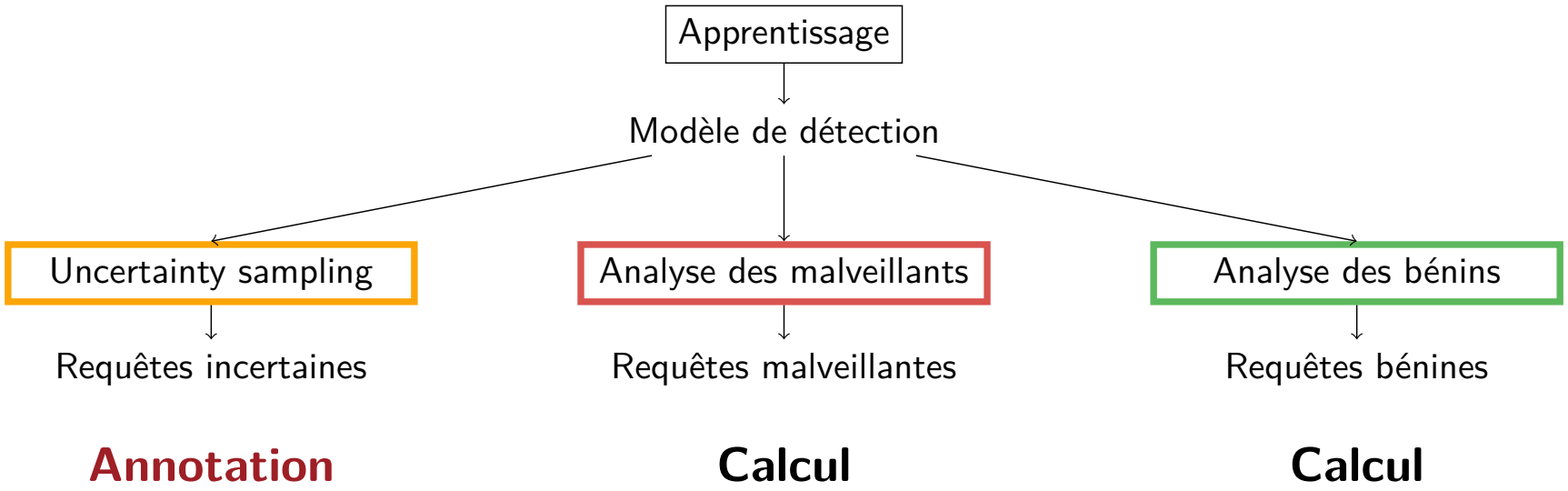
- Reduction de la complexité
- Annotations pendant les calculs



2- Réduire le temps d'attente

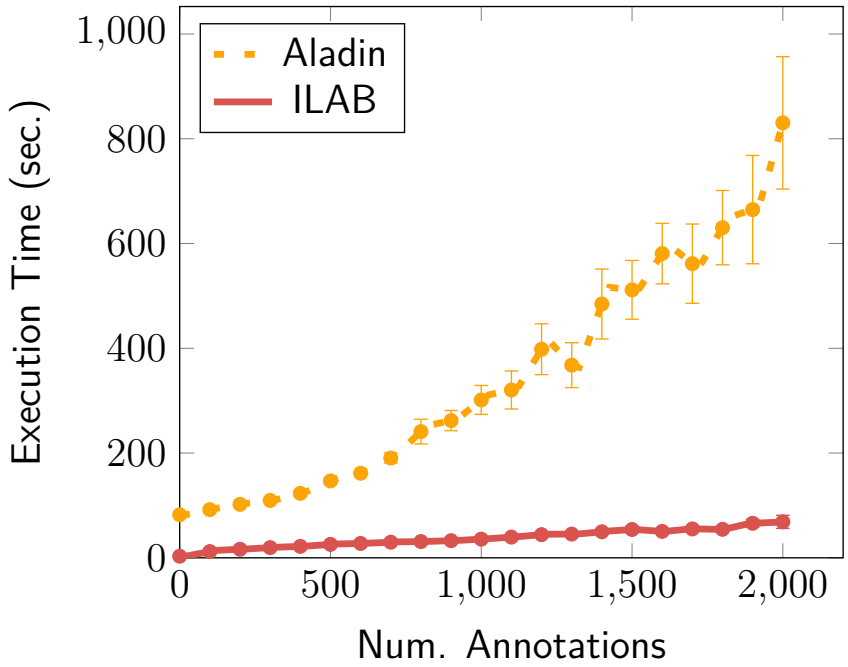
Diviser pour régner

- Reduction de la complexité
- Annotations pendant les calculs



2- Réduire le temps d'attente

Temps d'attente réduit grâce à ILAB

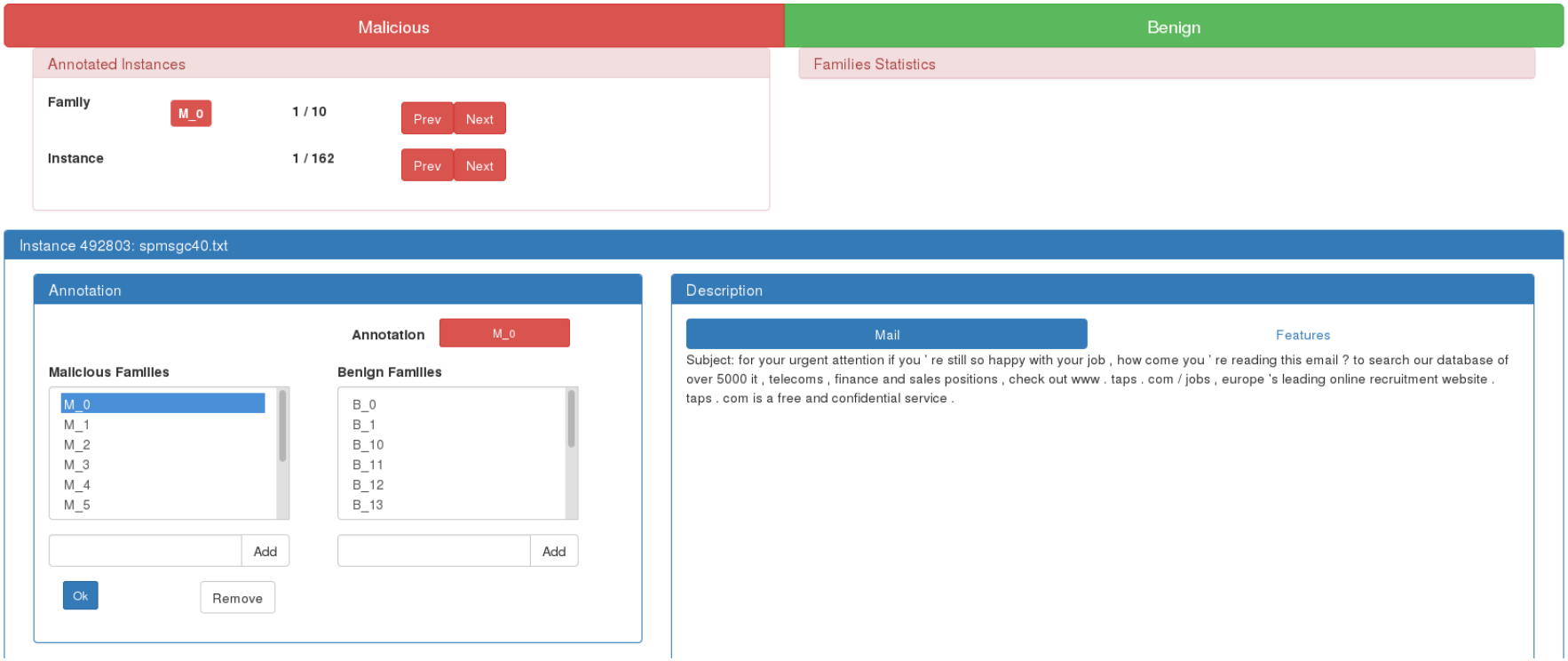


Aladin Stokes et al., Aladin: Active Learning of Anomalies to Detect Intrusions, 2008

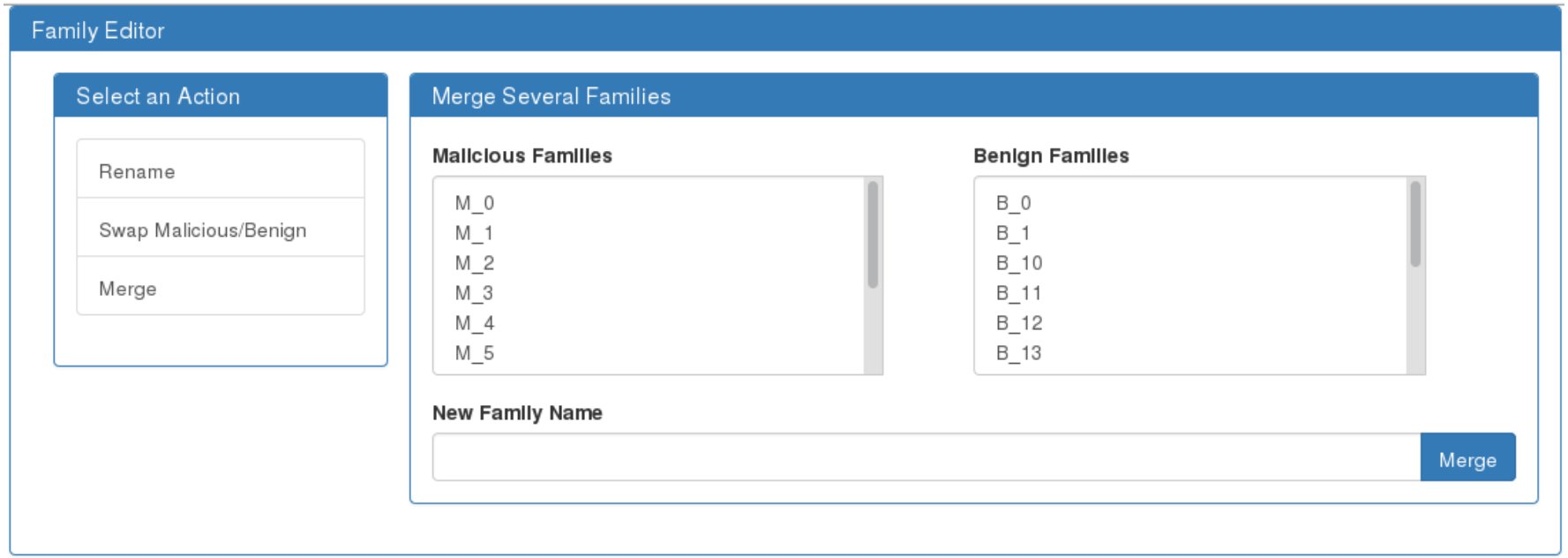
ILAB Beaugnon et al., ILAB: An Interactive Labelling Strategy for Intrusion Detection, RAID 2017

3- Interface utilisateur adaptée : Annotations

3- Interface utilisateur adaptée : Instances annotées



3- Interface utilisateur adaptée : Éditeur de familles



ILAB

Un système d'annotation adapté aux besoins des experts en sécurité

Répond aux défis

1. Détecter toutes les familles
2. Réduit le temps d'attente
3. Interface utilisateur adaptée

RAID 2017 Beaugnon et al., ILAB: An Interactive Labelling Strategy for Intrusion Detection

AICS 2018 Beaugnon et al., End-to-End Active Learning for Computer Security Experts

Machine Learning pour la détection d'intrusion

Bonnes pratiques

1. Un bon jeu de données annotées
2. Attributs discriminants
3. Modèle répondant aux contraintes opérationnelles
4. Validation du modèle

https://github.com/ANSSI-FR/SecuML

Machine Learning pour la détection d'intrusion

Bonnes pratiques

1. Un bon jeu de données annotées
2. Attributs discriminants
3. Modèle répondant aux contraintes opérationnelles
4. Validation du modèle

https://github.com/ANSSI-FR/SecuML