

Présentation

Apprentissage Sécurisé

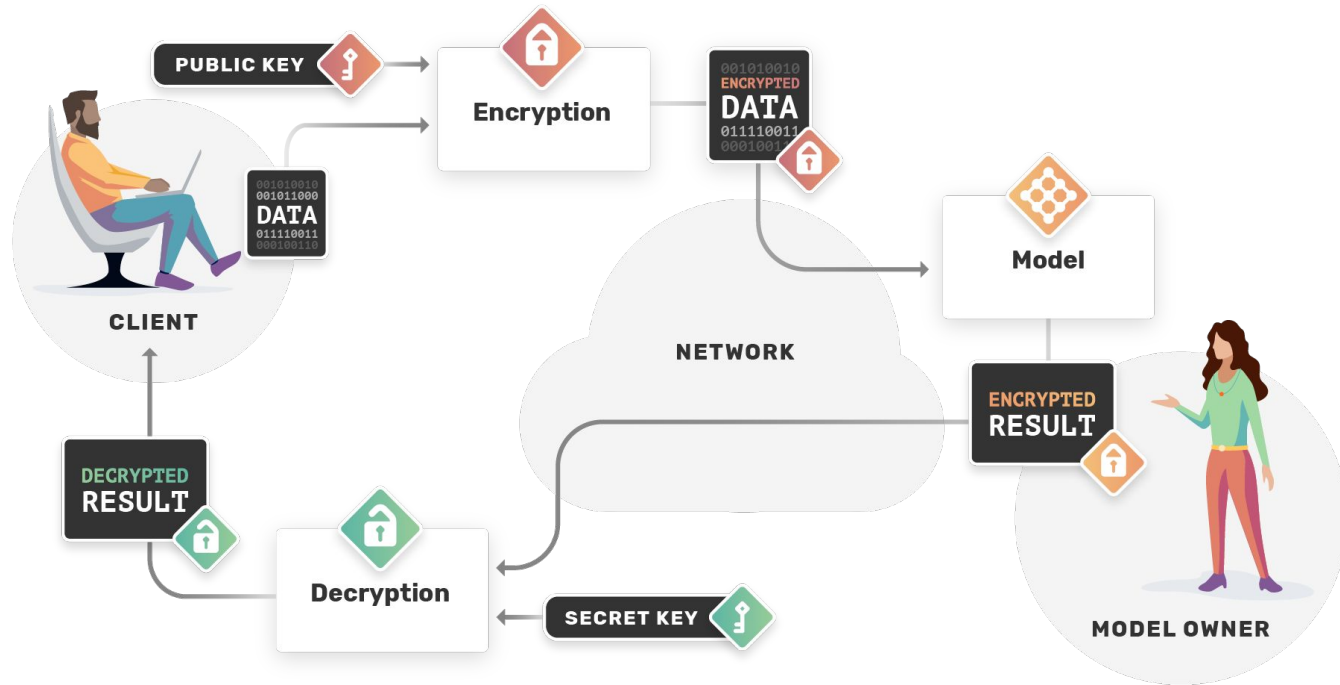


L'École des Ingénieurs Scientifiques

Sommaire

- **Introduction**
 - Contexte
 - Objectifs du projet
- **Méthodologie**
 - Organisation du projet
 - Etat de l'art

- **Développement réalisé**
 - Base de données
 - Type d'apprentissage
 - Type de chiffrement
 - Modèles d'apprentissage
 - Démo web
- **Bilan**
 - Nos difficultés
 - Conclusion



INTRODUCTION

- Contexte
- Présentation du projet

Contexte



Scénario 1 : Lutte contre la Criminalité

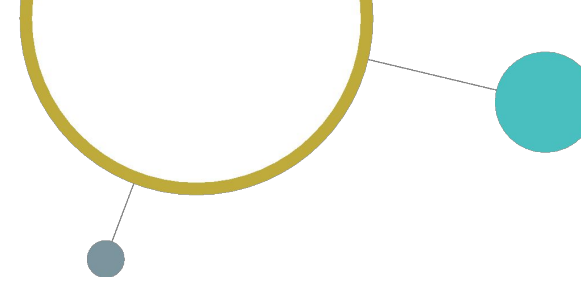
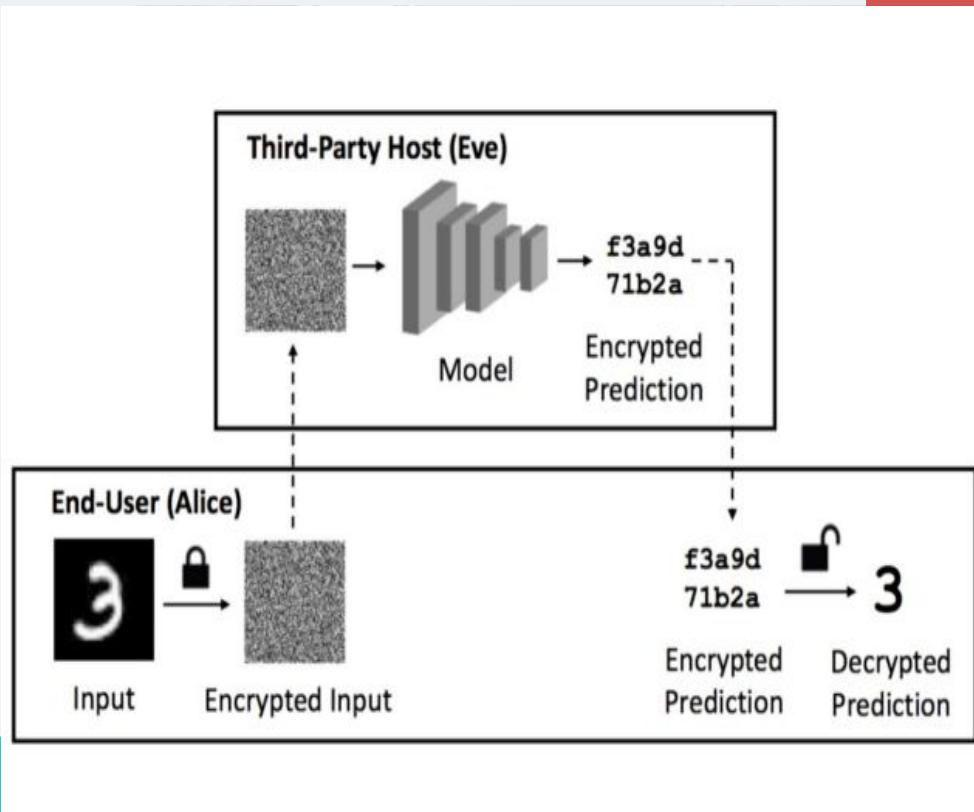
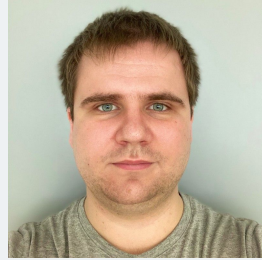
- Détection de contenu illégal chiffré grâce à l'IA
- Prévention de l'exposition directe des enquêteurs
- Maintien de la confidentialité des données et permettre l'identification d'activités illégales.



Scénario 2 : Confidentialité Médicale

- Données chiffrées pour préserver la confidentialité des patients.
- Fournir des résultats sans accéder aux données sensibles.
- Chiffrement de données entre l'IA et le médecin grâce aux clés publiques/privées

Présentation du projet proposé par le département SAFE, du laboratoire GREYC



Développement d'une IA sur des données chiffrées

- Deux questions principales:
 - Peut-on réaliser des tâches de prédiction à partir de données protégées, sans compromettre leur intégrité ?
 - Quel est l'impact sur la performance (précision et temps de calcul) ?

- **Effectuer un état de l'art scientifique**
 - Synthèse des travaux de recherche
 - Identification des méthodes et techniques
- **Création de nos Bases de Données**
 - Animaux
- **Développer un modèle de prédiction**
- **Évaluer les performances**
 - Précision
 - Temps
 - Mémoire
- **Effectuer une démonstration web**



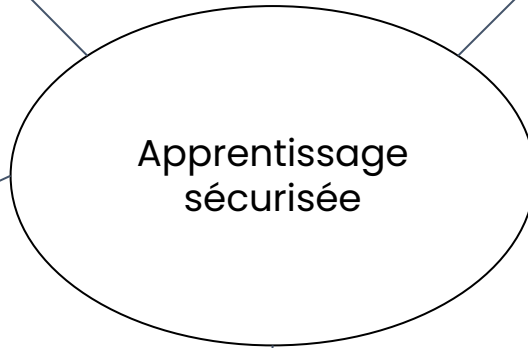
Objectifs du projet

Organisation du projet



Serveur Discord avec les tuteurs

Serveur Discord privé



Utilisation de Gitlab



Gestion de projet à jalon
Méthodologie Waterfall
(Cascade)



Réunion avec les tuteurs



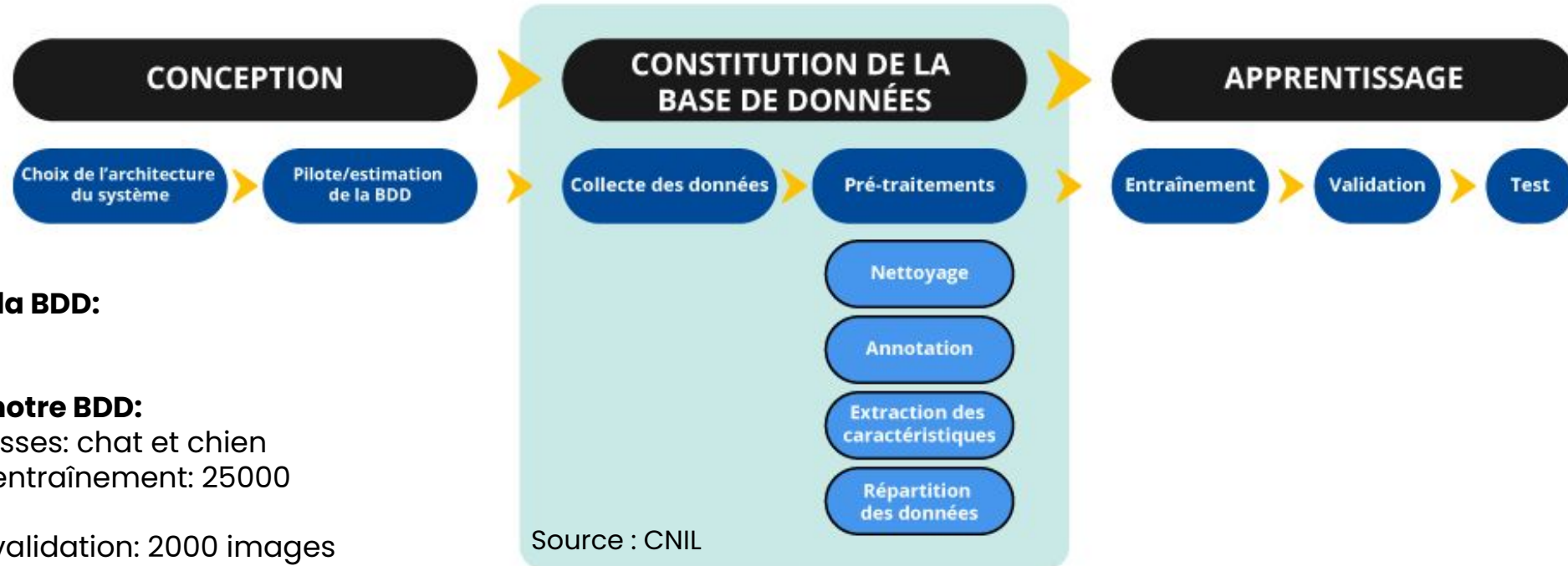
Etat de l'art

Article	Année	Principes	Base de données	Précision
CryptoNets: Apply-ing Neural Networks to Encrypted Data with High Throughput and Accuracy	2016	Leveled Homomorphic Encryption and Neural Networks	MNIST database	99% accuracy and around 59000 predictions per hour
Privacy Preserving Training and Evaluation with Homomorphic Encryption	2021	Implementation and test machine learning algorithms, including Logistic Regression(LR), Fully Connected Neural Network(FCNN), and Convolutional Neural Network(CNN).	MNIST database	about 98% accuracy with EncFCNN on 50 images and 100% with EncCNN on 100 images
Privacy-Preserving Classification on Deep Neural Network	2017	Application of secure computation in the context of machine learning	MNIST database	99.59% accuracy

Solutions de Modèle de réseaux neurones:

- **Fully Connected Neural Network**
- **Convolutional Neural Network**
- **IBMFHE**

Bases de données



Obtention de la BDD:

- Kaggle

Structure de notre BDD:

- Deux classes: chat et chien
- Dossier entraînement: 25000 images
- Dossier validation: 2000 images

Entraînement : Les images sont transformées (rotation, zoom,...) et sont labellisées (chat ou chien)

Validation : L'IA est évaluée sur une nouvelle base de données sans les transformations



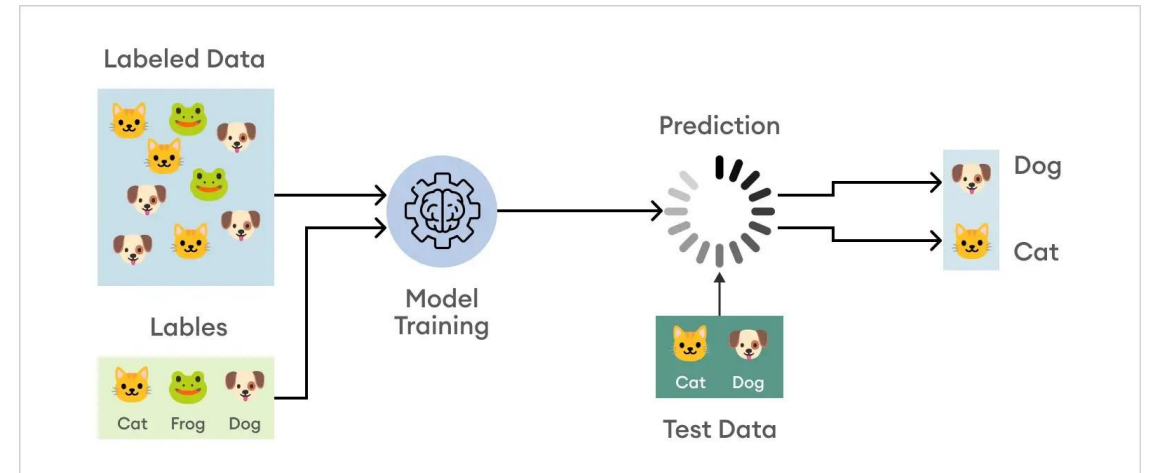
Prétraitement



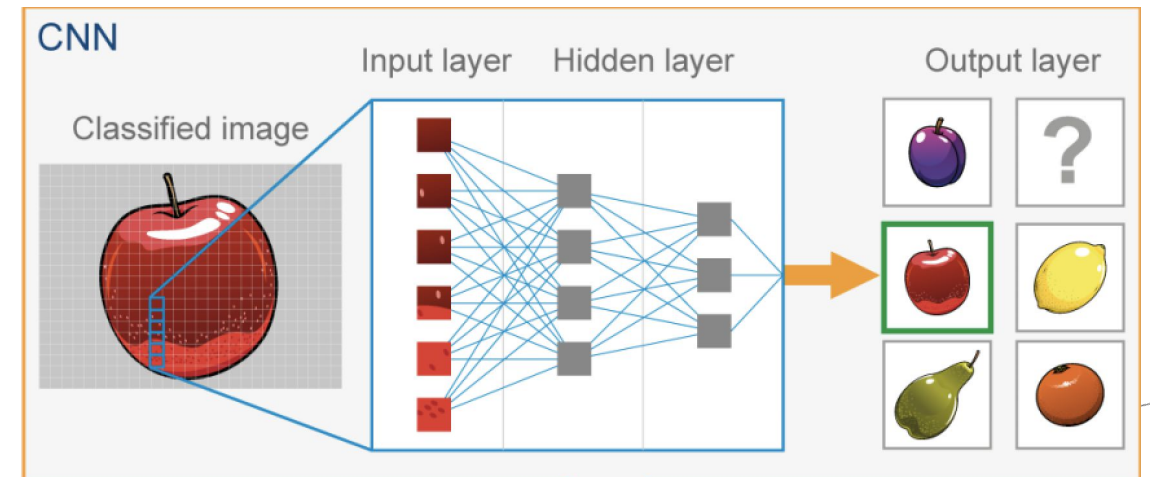
Chat

Apprentissage sur données en clair

- Apprentissage supervisé
- Préparation des données
- Construction et entraînement du modèle
- Validation et test du modèle
- Prédiction et interprétation des résultats



source : SuperAnnotate



source : Analytics Vidhya

Rappel de nos objectifs et organisation

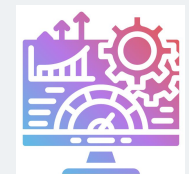
Objectifs

- Exploiter plusieurs techniques de chiffrement et d'implémentation de modèle d'apprentissage
- Identifier les modèles les plus performants (précision, temps, mémoire)
- Ajustement de notre modèle afin d'avoir une prédiction chiffrée (scénario 2)




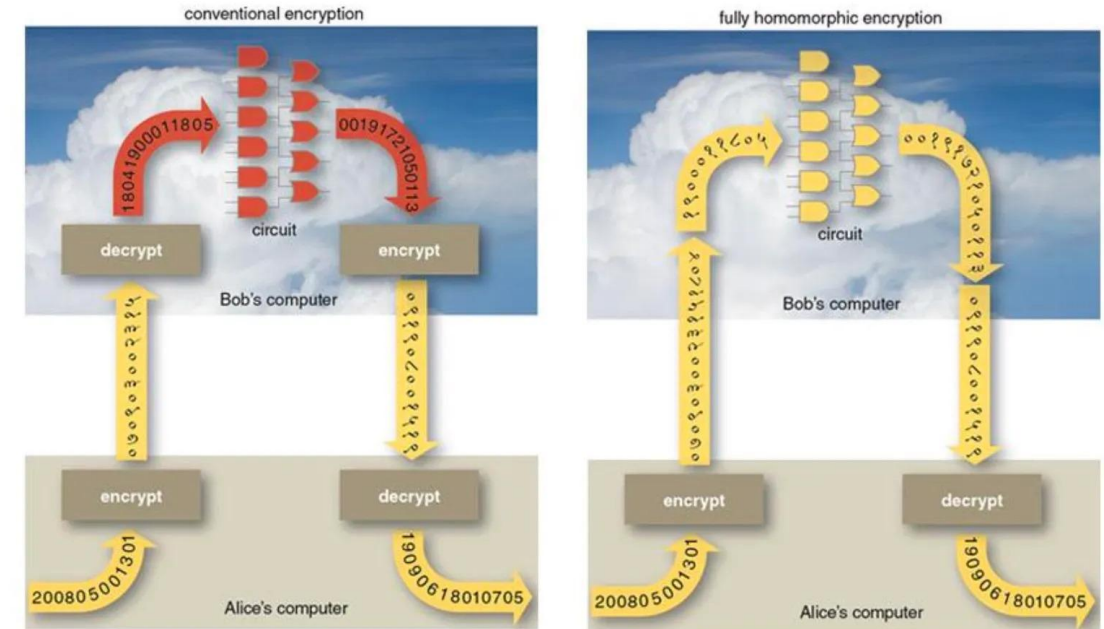
Protocole

- Base de données commune à chaque modèle
- Adaptation de modèles existant pour répondre à notre problématique
- Evaluation des performances sur chaque modèle
- Comparaison des résultats
- Optimisation des performances



Types de chiffrement

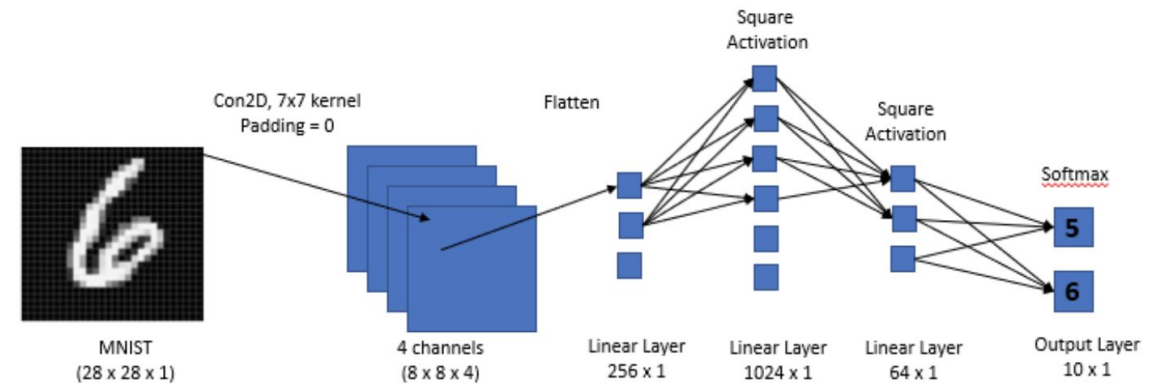
- Bibliothèque TENSEAL 
 - > Accès à une large gamme d'opérations
- **Chiffrement homomorphe : schéma CKKS**
 - > Permet d'effectuer des calculs sur des données chiffrées sans les déchiffrer, préservant ainsi la confidentialité.
 - Avantage : **Performance**
 - Inconvénient : **Limitation des opérations**
- **Fully Homomorphic Encryption (FHE)**
 - > Permet l'évaluation de fonctions arbitraires sur données chiffrées. Contrairement au CKKS, FHE n'est pas limité en terme d'opérateurs.
 - Avantage : **Expressivité (plus d'opérations)**
 - Inconvénient : **Complexité et performance**



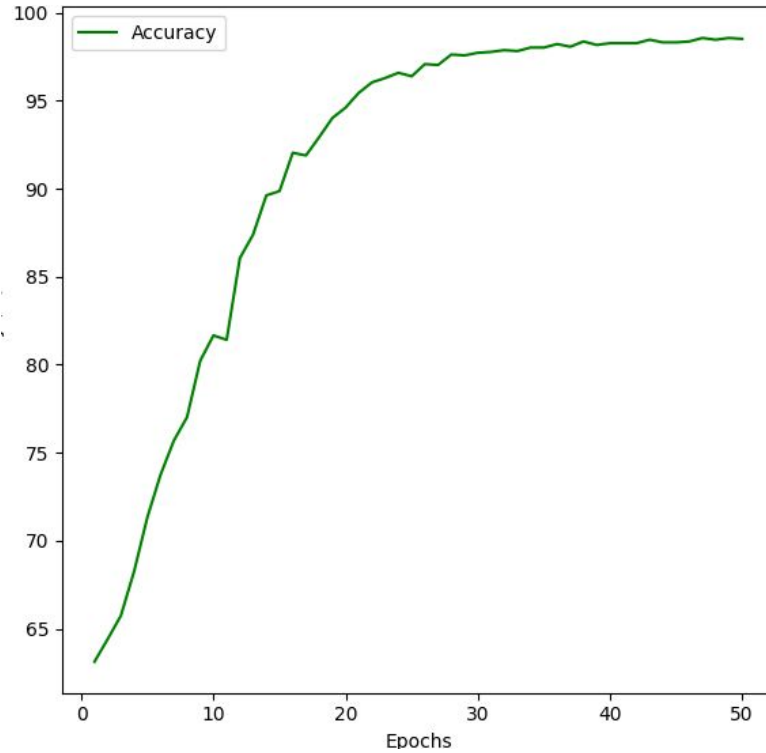
Source : Brian Hayes, *American Scientist* (www.americanscientist.org), septembre 2012

Modèle CNN encrypted

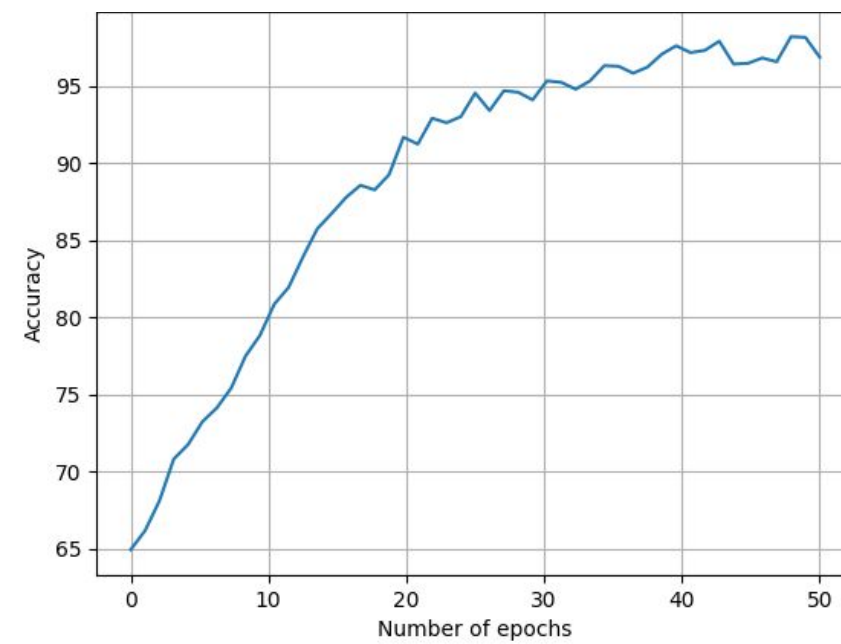
- **Modèle EncCNN basé sur CNN**
 -> Développement manuel d'un CNN
- **Chiffrement des poids avec CKKS**
- **Evaluation homomorphe**
 -> Prédiction sur des données chiffrées
- **Déchiffrement et interprétation**



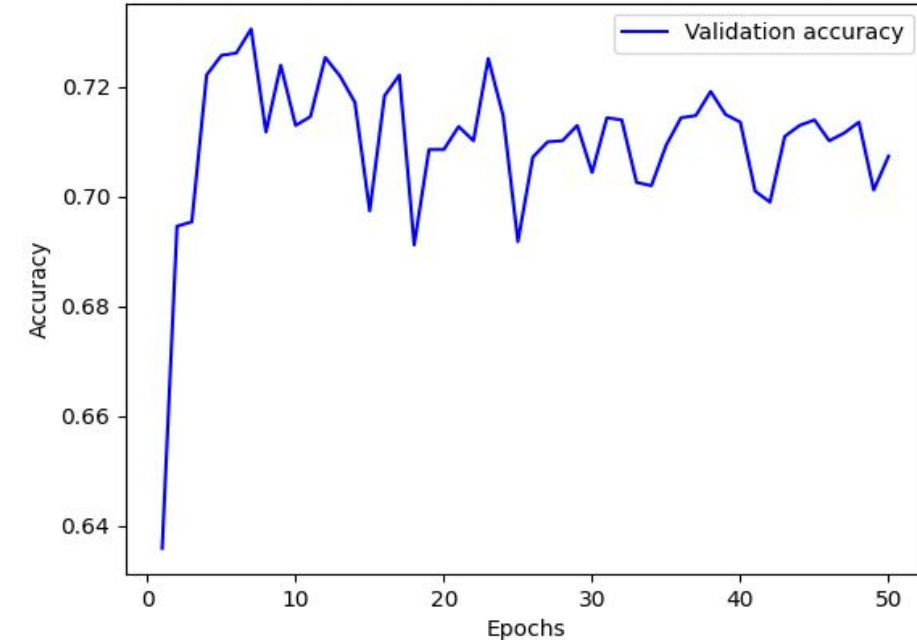
Comparaison des performances des différents modèles



Accuracy over epoch of FCNN model



Accuracy over epoch of CNN model



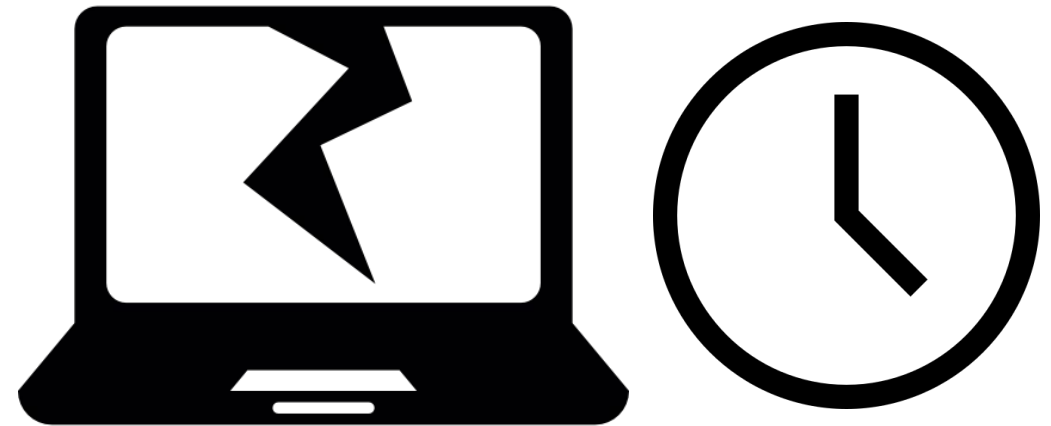
Accuracy over epoch of HElayer model

Model	CNN	EncCNN	FCNN	EncFCNN	HElayer
Accuracy	97%	97%	98%	50%	~70%
Prediction Time	0.43 s	2.4 s	0.41 s	19 s	0,48 s





Options d'optimisation pour le CNN

1. **Réglage Hyperparamètres**
2. **Optimisation de l'Architecture**
3. **Ajustement de la Précision du Chiffrement**
4. **Optimisation des Opérations Homomorphe**

Difficulté:
Problème matériels

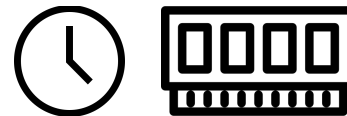


Nos difficultés

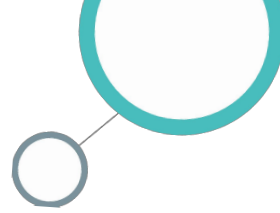
- Compréhension du sujet difficile au démarrage 
- Ordinateurs pas assez puissants, limités. 
- Compréhension des Gits difficile à cause du manque de simplifications/explications dans la doc. 
- Implémentation des différents modèles sur notre base de données. 

Objectifs non atteints:

- Évaluer nos performances en mémoire
- Optimiser notre modèle en temps et en mémoire



Conclusion



Points clés :

- Développement d'un modèle d'apprentissage chiffré
- Réalisation d'une démonstration web

Bilan :

- Satisfait de notre projet
- Communication très importante
- Complémentarité des deux filières
- Mauvaise estimation du matériel nécessaire

Perspective :

- Développement de nouveaux modèles associés à de nouveaux chiffrements

Question: Est-ce viable sur des données illégales/confidentielles même si le temps est long ?



Encrypted

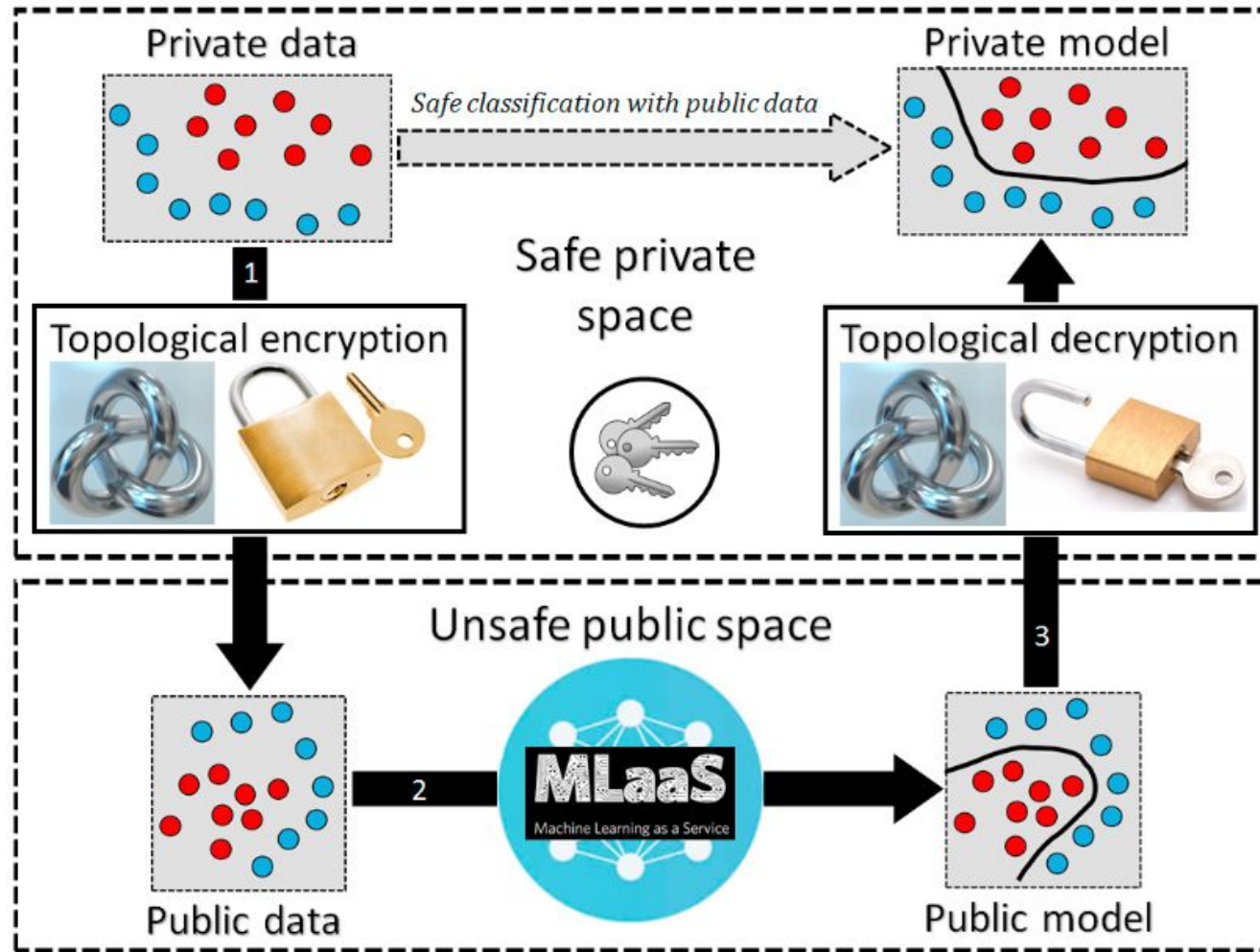
MERCI
pour votre écoute



L'École des Ingénieurs Scientifiques

ANNEXES

Développement d'un modèle d'apprentissage sur données chiffrées



Démonstration Web

Objectif: Visualisation pédagogique de notre projet

Mise en place de la démo web : Bibliothèque streamlit de Python

Fonctionnement: Choix d'une image, choix chiffrement, choix du modèle et prédiction.

Démo de détection de chat ou de chien sur des images chiffrées

Choisir une image de chat ou de chien



Drag and drop file here

Limit 200MB per file • JPG, PNG, JPEG

Browse files

Choisir le chiffrement

ENcCNN

Prédire avec ENcCNN

C'est probablement un chat 🐱

Probabilité d'être un chat

100.00%

Probabilité d'être un chien

0.00%

Chat 100.00%

Chien
0.00%

Temps d'exécution de la prédiction en secondes : 2.5