Présenté par nous :

Test plus amples sur les MLP, a la fois bi-classe multiclasse.
Bons résultats pour le bi-classe même pour un faible entrainement.
Test également avec 4 classes d'attaques : très bons résultats
but sprint 3 : affiner les modèles.

Types d'attaques implémentés :

Attaques FGSM sur les modèles pytorches. Test pas encore fini mais mis en place.
Attaques FGSM sur les randoms forest : le MLP de substitution est perturbé très éfficacement mais le RF pas réellement
Attaque HSJ et FSGM sur knn : très bon résultat pour le HSJ (perte de 99% de précision).

Client en déplacement pendant les 3 prochains jours : parler des résultats par mails
Attaques de type HSJ : boite noire accès qu'aux entrée et sorties du modéle.
Client : pendant soutenance explication du fonctionnement des attaques.

Attaques knn : très faible à l'ataque HSJ (baisse à 1%) mais pas FGSM

Client:

Client satisfait des résultats, penser a expliquer le principe de chaque attaque.
Analyser les features considéré : modification possible , range de la modification.
Pour la présentation penser a prendre une petite partie (4-5 ligne et regarder l'évolution de leur détection au fur et a mesure des attaques adverse).
Faire un dossier test général pour la démo.