Université du Luxembourg
snt@uni.lu - snt.uni.lu

Adversarial Attacks in Machine Learning/Deep Learning

Votre rôle

This internship would be dedicated to one of the following lines of research:

- (adversarial ML) improving our understanding of adversarial examples. Despite achieving superhuman accuracies, state-of-art neural networks in computer vision and natural language processing are easily vulnerable to adversarial examples, i.e., small imperceptible perturbations on inputs that lead to large changes in outputs. Despite many papers written on the field, the causes of adversarial examples are not well understood. Your objective would be to identify new properties of adversarial examples that would improve our understanding of such vulnerabilities and/or that proves useful to design stronger attacks. Practical applications would be either on classification for computer vision or natural language processing.
- (adversarial ML) adversarial examples in the presence of distributional shift. In the literature of adversarial ML, many attacks assume the access to either the data used to train the targeted model or to some data generated by the same data distribution. Some prior work show that in some specific settings, it is possible to craft adversarial examples on a model trained on a different data distribution than the data used to train the model we wish to target. However, the experimental settings are quite narrow, and this direction remains mostly unexplored.
- (natural generalization of neural networks) finding permutation symmetries for neural networks. Despite being highly multi-modal, previous work shows that it is possible to find paths in the space of neural network weights (the parameter space) that connect two independently trained neural networks where each model along the path have similar performances. These observations have powerful implications to understand why neural networks work and can be easily trained with stochastic gradient descent. However, finding those permutations is a hard problem, and remains open in many experimental settings.

Keywords :

deep learning, adversarial machine learning, adversarial examples, representation learning, generalization gap, ensemble techniques, geometrical interpretation of neural networks

Vous serez amené à effectuer les tâches suivantes :

- Passer en revue la littérature scientifique et dresser un état de l'art

- Implémenter divers algorithmes
- Mettre en œuvre une stratégie expérimentale
- Analyser les résultats de vos expériences
- Communiquer avec des doctorants et chercheurs de l'université pour rendre compte de vos résultats

L'équipe dans laquelle vous travaillerez

- Martin Gubri
- Dr. Maxime Cordy: Superviseur
- Prof. Yves Le Traon: Directeur de l'équipe de recherche SerVal

Votre profil

- Étudiant.e Bac +5 en école d'informatique, université ou école d'ingénieur avec une formation en statistiques, machine learning et analyse de données
- Disposant d'une 1ère expérience sur des projets de machine learning et/ou deep learning
- Bonne maîtrise des langages de programmation (Python / R) ainsi que des frameworks en traitement de données (Pandas, Numpy), en visualisation de données (Matplotlib, Seaborn) et en Machine Learning/Deep Learning (Scikit-learn, Tensorflow, Keras, Torch)
- Curieux.euse, agile et possédant de bonnes capacités d'analyse et de synthèse
- Doté.e d'un bon relationnel et d'un fort esprit d'équipe
- Vous disposez d'un niveau de compréhension et d'expression en anglais vous permettant de communiquer avec des doctorants et chercheurs venu du monde entier.

Ce qui vous attend au SnT...

Des infrastructures passionnantes et des laboratoires uniques. Sur les deux campus du SnT, nos chercheurs peuvent se promener sur la lune au LunaLab, construire un nanosatellite ou contribuer à améliorer les véhicules autonomes. Les chercheurs du SnT s'engagent dans des projets axés sur la demande. Grâce à notre programme de partenariat, nous travaillons sur des projets avec plus de 45 partenaires industriels.

Faites partie d'une famille multiculturelle. Au SnT, nous comptons plus de 60 nationalités. Tout au long de l'année, nous organisons des événements de renforcement de l'esprit d'équipe, des activités de mise en réseau, etc.

Université du Luxembourg
snt@uni.lu - snt.uni.lu

En résumé

- Type de contrat : Stage 4 à 6 mois
- Début du stage : Selon disponibilité
- Temps de travail: Plein temps 40.0 heures par semaines
- Location: Luxembourg
- Gratification : ~ 1200€ mensuel

Comment postuler ?

Les candidatures doivent comprendre :

- CV
- Lettre de motivation

Merci d'envoyer ces documents à :

- martin.gubri@uni.lu

À propos de l'université du Luxembourg...

L'Iniversité du Luxembourg cherche à recruter des chercheurs au SnT (Interdisciplinary Centre for Security, Reliability and Trust).

Le SnT mène des recherches interdisciplinaires sur les systèmes et services ICT (Information and Communication Technologies) sûrs, fiables et dignes de confiance, souvent en collaboration avec des partenaires industriels, gouvernementaux ou internationaux. Le SnT est actif dans plusieurs projets de recherche internationaux financés par le programme Horizon2020 et l'Agence spatiale européenne. Pour plus d'informations, vous pouvez consulter : https://wwwfr.uni.lu/snt