



Master Informatique

Proposition de sujet de Stage Recherche 2018-2019

Titre : Analyse des flux continus de données par les techniques de Machine Learning

Laboratoire : LIAS

Encadrant(s) : Prof. Allel HADJALI (allel.hadjali@ensma.fr)
Dr. Stéphane JEAN (stephane.jean@ensma.fr)
Dr. Mickael BARON (mickael.baron@ensma.fr)

Mots clés : Flux de données, analyse de données, capteurs, machine learning

Sujet : L'Internet of Thing (IoT) est un nouveau paradigme référant à un ensemble d'objets connectés via internet pour fonctionner d'une manière collaborative. En 2020, le nombre total d'objets interconnectés utilisés est estimé entre 25 à 50 milliards d'objets. Comme un objet connecté est synonyme de « collecte et traitement des données en temps réel », IoT deviendra ainsi parmi les plus grandes et majeures sources productrices de données dans les prochaines années. Cette technologie a reçu et continue de recevoir un intérêt croissant dans les applications du monde réel issues de l'industrie (aéronautique), de transport, de la santé, de la vie quotidienne des personnes, etc.

Les données capteurs IoT sont générés sous formes de flux. En plus de leur grande volumétrie, ces données sont intrinsèquement incertaines, erronées et bruitées d'une part, et de différents types, distribuées et continues d'autre part. Avec cette variation dans la qualité des données IoT, et afin de tirer profit et donner du sens à ces données, un des défis à relever par l'industrie (ou les personnes) est de faire migrer les modèles d'analyse de données traditionnels vers des modèles d'analyse modernes et intelligents où le concept de données intelligentes constitue un bon représentatif des données IoT. Ces modèles, fondés sur des techniques de machine learning et de l'intelligence artificielle, permettent par exemple de prédire des résultats désirés à partir des données collectées.

Ce travail de Master constitue les premières réflexions sur une approche, guidée par les techniques de Machine Learning, pour une analyse pertinente des flux continus de données provenant de multi-capteurs. Il comprendra notamment les points suivants :

- (i) Réexaminer les caractéristiques des données IoT dans le monde réel.
- (ii) Etudier comment les algorithmes de machine learning peuvent être appliqués sur les données IoT.

Une étude de cas issus de l'industrie sera considérée pour illustrer et valider nos propositions.



Master Informatique

Biblio :

- Aimad Karkouch, [Hajar Mousannif](#), [Hassan Al Moatassime](#), [Thomas Noël](#): Data quality in internet of things: A state-of-the-art survey. [J. Network and Computer Applications 73](#): 57-81 (2016)
- [Zheng Yan](#), [Peng Zhang](#), Athanasios V. Vasilakos: A survey on trust management for Internet of Things. [J. Network and Computer Applications 42](#): 120-134 (2014)
- P. Barnaghi, M. Bermudez-Edo, R. Tonjes, Challenges for quality of data in smart cities, Journal of Data and Information Quality (JDIQ) 6 (2-3) (2015) 6.
- D. Toshniwal, et al., Clustering techniques for streaming data-a survey, in: Advance Computing Conference (IACC), 2013 IEEE 3rd International, IEEE, 2013, pp. 951–956.
- T.M. Mitchell, Machine Learning, cGraw – Hill International, 1997.
- Taiwo Oladipupo Ayodele, Types of Machine Learning Algorithms, New Advances In Machine Learning, Yagang Zhang (Ed.), InTech, 2010
- Barber, D., Bayesian Reasoning and Machine Learning, Cambridge University Press, 2012

Lieu du stage : LIAS, ENSMA

Parcours conseillé : Données

UEs optionnelles conseillées : IDD, AFGL

Etudiant pressenti (éventuellement) : O. Sadouni