

Offre de stage

Sujet : Analyse automatique de signaux sociaux non verbaux dans conversations téléphoniques.

Encadrement

Giovanna Varni

Lieu et dates du stage

Telecom ParisTech, 46 rue Barrault, 75013 Paris

Date de début du stage : février 2018

Équipe(s) d'accueil de la thèse

Département IDS, équipe S2A, thème de recherche : Analyse automatique des données sociales

Mots clés

Analyse des signaux sociaux, interaction humain-humain, machine learning,

Sujet détaillé

L'analyse des signaux sociaux a comme but le développement de machines douées d'une intelligence sociale et émotionnelle [1]. Ce nouveau domaine de recherche est en plein essor et a permis aux informaticiens d'aborder de plus en plus la théorie et les modèles de la communication verbale (ce qui est dit) et non verbale (comment ce qui est dit/fait est exprimé) qui ont été amplement formulés et étudiés en psychologie [2]. Toutefois, à ce jour, le rôle joué par la communication non verbale dans les conversations téléphoniques n'est pas encore clair. Le sujet de stage porte sur le développement de méthodes d'analyse de signaux sociaux non verbaux dans un corpus de conversations téléphoniques [3]. En particulier, le focus du stage sera sur la dominance qui est un signal social très important dans l'échange d'informations entre deux ou plusieurs personnes. La dominance est un trait personnel durable qui exprime le tempérament et les prédispositions comportementales d'une personne pendant l'interaction [4]. Le stagiaire travaillera sur un ensemble de caractéristiques non verbales identifiées en psychologie comme signes de dominance dans l'interaction ; et construira des modèles computationnels de dominance en s'appuyant sur différentes méthodes d'analyse (par exemple sequence mining) .

Ce stage contribuera à améliorer la palette de méthodes et de modèles computationnels appropriés pour les technologies conversationnelles. A titre d'exemple, ces méthodes et modèles pourront aider

à découvrir les relations hiérarchiques intervenantes dans deux ou plusieurs personnes dans scénarios différents.

Ce stage se déroulera dans le cadre d'une collaboration avec l'Université de Glasgow.

La Chaire Machine Learning for Big Data

Le traitement statistique des masses de données convoque à la fois mathématiques appliquées et informatique, à travers une discipline en plein essor : le Machine Learning ou apprentissage statistique.

Créée en septembre 2013 avec le soutien de la Fondation Télécom et financée à hauteur de près de 2 M€ par quatre entreprises partenaires : Criteo, PSA Peugeot Citroën, Safran et BNP Paribas, la Chaire Machine Learning for Big Data est portée par le mathématicien Stéphan Cléménçon, Enseignant-Chercheur, Professeur au sein du Département du Traitement du Signal et des Images à Télécom ParisTech.



Proposant cinq axes de recherche méthodologiques, enrichis par des applications industrielles concrètes, cette Chaire a pour objectif d'animer, en interaction avec ses partenaires, une activité de recherche de pointe en Machine Learning, ainsi que de proposer des programmes de formation.

La variété des données aujourd'hui disponibles (nombres, images, textes, signaux), leur grande dimension et leur volumétrie rendent souvent inopérantes les méthodes statistiques traditionnelles reposant sur le prétraitement humain et un long travail de modélisation. Le Machine Learning vise donc à élaborer et étudier des algorithmes, à vocation prédictive le plus souvent, permettant à des machines d'apprendre automatiquement à partir des données et à effectuer des tâches de façon performante.

Les avancées technologiques, l'omniprésence des capteurs (systèmes embarqués, objets connectés, Internet...) et l'explosion des réseaux sociaux s'accompagnent d'un véritable déluge de données, propulsant les sciences de l'information au centre du processus de valorisation des masses de données. En plus de la collecte et du stockage, l'enjeu est de pouvoir analyser ces données afin d'optimiser les décisions et mettre au point de nouvelles applications.

Au-delà du buzz médiatique dont il fait l'objet, le Big Data est donc un sujet stratégique majeur, au cœur d'enjeux économiques et sociétaux considérables. Son impact est désormais perçu dans presque tous les secteurs de l'activité humaine : de la recherche scientifique à la médecine en passant, entre autres, par la finance, le bâtiment, l'e-commerce, la défense ou les transports.

En savoir plus sur la Chaire, ses axes de recherche, ses activités, ses publications :

<http://machinelearningforbigdata.telecom-paristech.fr>

Profil du candidat

Etudiant de master 2 recherche

- Apprentissage statistique / reconnaissance des formes
- Traitement de signaux
- Bon niveau en programmation (Python)
- Bon niveau d'anglais

Candidatures

à envoyer à giovanna.varni@telecom-paristech.fr:

- Curriculum Vitae
- Lettre de motivation personnalisée expliquant l'intérêt du candidat sur le sujet
- Relevés de notes des années précédentes
- Contact d'une personne de référence

Les candidatures incomplètes ne seront pas examinées.

Référence

[1] Vinciarelli, A., Pantic, M. Bourlard, H. "Social Signal Processing". Image and Vision Computing, 27 (12), 2009.

[2] Knapp, M. L., Hall, J. A. Nonverbal communication in human interaction. Thomson Learning.

[3] Polychroniou, A., Salamin, H., Vinciarelli, A. "The SSPNet mobile corpus: social signal processing over mobile phones," in Proceedings Language Resources and Evaluation Conference, 1492–1498, 2014.

[4] Ridgeway, C. L., "Nonverbal behavior, dominance, and the basis of status in task groups", American Sociological Review, 52, 683–694, 1987.