



Offre de stage

Prédiction multi-labels de compétences sociales dans des vidéos d'entretiens d'embauche par des réseaux de neurones récurrents

A propos de la chaire

La Chaire Data Science and Artificial Intelligence for Digitalized Industry and Services, portée par Florence d'Alché-Buc, enseignante-chercheuse dans le département Image, Données, Signal de Télécom Paris, la chaire DSAIDIS réunit cinq partenaires industriels : Airbus Defence & Space, Engie, Idemia, Safran et Valeo. Son objectif général est de développer, en liaison étroite avec les partenaires, une formation et une recherche de niveau international.

Ses quatre principaux axes de recherche sont :

1. Analyse et prévision de séries temporelles (Predictive Analytics on Time Series) ;
2. Exploitation de données hétérogènes, massives et partiellement étiquetées (Exploiting Large Scale and Heterogeneous, Partially Labelled Data) ;
3. Apprentissage pour une prise de décision robuste et fiable (Learning for Trusted and Robust Decision) ;
4. Apprentissage dans un environnement dynamique (Learning through Interactions with a Changing Environment).

Plus d'information sur la Chaire : www.telecom-paris.fr/dsaidis

Description du stage

Encadrement

Chloé Clavel, Léo Hemamou

Lieu et dates du stage

Adresse : Télécom Paris, 19 Place Marguerite Perey, 91120 Palaiseau

Date de début du stage : Premier trimestre 2020

Équipe(s) d'accueil de la thèse

Département IDS, équipe Signal, Statistiques et Apprentissage (S2A)

Mots clés

prédiction multi-labels, réseaux de neurones récurrents, analyse de compétences sociales, entretien d'embauche, social computing

Sujet détaillé

Le développement des nouvelles technologies impacte tous les secteurs d'activités, y compris celui des Ressources Humaines, que ce soit durant la recherche de candidats ou dans le processus de sélection. Ainsi, l'entretien vidéo différé permet d'organiser en asynchrone des entretiens avec des candidats et de les évaluer. Les candidats se connectent à une plateforme, se filment pendant qu'ils répondent à des questions définies a posteriori par les recruteurs. La plateforme permet ensuite à plusieurs recruteurs d'évaluer le candidat, d'échanger entre eux et d'inviter éventuellement le candidat à un entretien en face à face. Les recruteurs établissent au préalable un questionnaire de recrutement et y associent des critères d'évaluation. Le/la candidat-e reçoit une invitation pour répondre à ces questions et enregistre en vidéo ses réponses selon ses disponibilités dans un temps limité. Il ou elle n'a pas connaissance des questions à l'avance et ne peut pas se réenregistrer afin de préserver la spontanéité de ses réponses. Le recruteur reçoit les vidéos sur une interface et peut ainsi comparer et évaluer les différents profils avec ses équipes selon les critères définis précédemment. De plus en plus d'entreprises font le choix de ce type d'entretien comme outil de présélection. Le choix d'un tel outil est motivé par l'accès à un plus grand nombre et une plus grande diversité de candidats et par la réduction du temps de traitement et de prise de rendez-vous [Torres2017].

Le nombre de candidatures vidéo devient difficile à traiter « manuellement » par un ou deux recruteurs. Il devient donc nécessaire d'envisager une aide pour le recruteur devant traiter parfois plusieurs dizaines (voire centaines) d'entretiens vidéo. De plus, le développement d'une telle aide pourra aussi permettre aux candidats de s'entraîner à l'exercice de l'entretien vidéo différé grâce à une évaluation automatique.

Le développement de modèles de comportements non verbaux et l'étude de leur influence dans les entretiens d'embauche suscitent un engouement récent [Hoque2013][Naim2018][Rasipuram2018][Hemamou2019b].

Un défi important pour la conception de ces modèles réside dans la transparence de ces modèles pour 1) construire une relation de confiance entre le recruteur et le système automatique 2) détecter possiblement des comportements non verbaux influents lors des entretiens d'embauche 3) permettre aux recruteurs de prendre conscience de ses processus d'évaluation.

Un premier modèle d'apprentissage profond multimodal [Hemamou2019a] a été proposé afin de répondre à ce défi.

Cependant, ce premier modèle ne fournit qu'une information globale quant à la performance du candidat. Or, de nombreux critères pouvant expliquer l'employabilité des candidats ont été mis en exergue[Huffcutt2001] tels que les compétences communicationnelles ou les traits de personnalité. De plus, lors de la construction d'un entretien d'embauche, un recruteur fournit une grille d'évaluation grâce à laquelle il va juger un candidat. Cette grille nous fournit des indications précises quant aux dimensions évaluées, néanmoins, ces critères peuvent être variables d'un recruteur à l'autre de part l'unicité des postes ou de part la culture d'entreprise [Salgado2017]

Ce stage viendra renforcer l'axe 2 (Exploitation de données partiellement étiquetées) de la chaire DSAIDIS dans le cadre des entretiens d'embauches différés. Il porte sur le développement d'approches basées sur des modèles multi-labels[Garcia2019] et d'apprentissage de représentation de labels (label embedding italique) [Pappas2019][Wang2018]. Partant des travaux de [Hemamou2019a], le stagiaire développera des architectures neuronales récurrentes qui seront pensées afin de résoudre des problèmes multi-labels ou de représentations de labels. Plus précisément, le stagiaire sera amené à :

- Etudier les compétences générales recherchées en entretien d'embauches différés par une méthode d'analyse de corpus,
- Etudier la littérature récente sur les modèles multi-labels, les méthodes de représentation de labels,
- Implémenter des réseaux de neurones récurrents basés sur des unités LSTM ou GRU pour la résolution de problème de multi-labels,
- Proposer et implémenter un modèle pour la résolution du problème multi-labels par une méthode de plongement de labels (label embedding).

Profil du candidat

- Etudiant titulaire d'un master 2 recherche
- Apprentissage statistique / reconnaissance des formes
- Traitement de la parole, traitement du langage naturel
- Bon niveau en programmation (Java, C/C++, Python)
- Bon niveau d'anglais

Candidatures

A envoyer en un seul document pdf à chloe.clavel@telecom-paristech.fr, leo.hemamou@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.

Références

- [Torres2017] Torres, E.N. et Mejia, C. 2017. Asynchronous video interviews in the hospitality industry: Considerations for virtual employee selection. *International Journal of Hospitality Management*. 61, (2017), 4-13. DOI:<https://doi.org/10.1016/j.ijhm.2016.10.012>.
- [Hemamou2019a] Hemamou, L. et al. 2019. HireNet : a Hierarchical Attention Model for the Automatic Analysis of Asynchronous Video Job Interviews. *AAAI* (2019).
- [Hemamou2019b] Hemamou, L., Felhi, G., Martin, J. et Clavel C., 2019. Slices of Attention in Asynchronous Video Job Interviews. (2019).
- [Naim2018] Naim, I. et al. 2018. Automated Analysis and Prediction of Job Interview Performance. *IEEE Transactions on Affective Computing*. 9, 2 (avr. 2018), 191-204. DOI:<https://doi.org/10.1109/TAFFC.2016.2614299>.
- [Rasipuram2018] Rasipuram, S. et Jayagopi, D.B. 2018. Automatic assessment of communication skill in interview-based interactions. (2018).
- [Huffcutt2001] Huffcutt, A.I. et al. 2001. Identification and meta-analytic assessment of psychological constructs measured in employment interviews. *Journal of Applied Psychology*. 86, 5 (2001), 897-913. DOI:<https://doi.org/10.1037//0021-9010.86.5.897>.
- [Salgado2017] Salgado, J.F. 2017. Personnel Selection. *Oxford Research Encyclopedia of Psychology*. Oxford University Press.
- [Pappas2019] Pappas, N. et Henderson, J. 2019. GILE: A Generalized Input-Label Embedding for Text Classification. *Transactions of the Association for Computational Linguistics*. 7, (2019), 139-155. DOI: https://doi.org/10.1162/tacl_a_00259.
- [Wang2018] Wang, G. et al. 2018. Joint embedding of words and labels for text classification. *ACL 2018 - 56th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference (Long Papers)*. 1, (2018), 2321-2331. DOI:<https://doi.org/10.18653/v1/p18-1216>.
- [Hoque2013] Hoque, M.E. Courgeon, M., Martin, J.C., Mutlu, B et Picard, R.W. 2013. MACH. *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing - UbiComp '13* (New York, New York, USA, 2013), 697.
- [Garcia2019] Alexandre Garcia, Pierre Colombo, Florence d'Alché-Buc, Slim Essid and Chloé Clavel, [From the Token to the Review: A Hierarchical Multimodal approach to Opinion Mining](#), *EMNLP 2019 - Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*