

Personnalisation de retours d'information émotionnelle par recommandation d'émotions

Offre de thèse - LIPN, Université Sorbonne Paris Nord, France
Encadrement : Davide Buscaldi (HDR), Gaël Guibon

Mai 2025

Titre du sujet : Personnalisation de retours d'information émotionnelle par recommandation d'émotions

- **Unité de recherche** : LIPN CNRS UMR 7030
- **Discipline** : Informatique
- **Direction de thèse** : Davide BUSCALDI ; co-encadrement : Gaël GUIBON
- **Contact** : buscaldi@lipn.univ-paris13.fr ; guibon@lipn.univ-paris13.fr
- **Domaine de recherche** : Traitement automatique du langage (TAL) ; Affective Computing
- **Mots clés** : recommandation d'émotions, personnalisation, emotional feedbacks (retours d'information émotionnelle), agent conversationnel

1 Contexte

Prédire et détecter des émotions ignore bien trop souvent l'aspect subjectif et personnel des états émotionnels. Chaque humain a une réponse émotionnelle différente selon un même contexte donné (Kochan, 2013), il est donc nécessaire pour un agent conversationnel d'offrir cette même différence d'émotion. De plus, il a été montré que le reflet émotionnel (Picard, 2000), aussi appelé "contagion" dans une de ses variantes (Brave and Nass, 2007), entraîne une perception plus crédible d'un agent conversationnel dans une relation humain-machine. Pour répondre à ce besoin, nous nous attellerons à recommander les émotions dans un contexte conversationnel de manière personnalisée par l'utilisateur humain. La recommandation servira à guider l'affinage du grand modèle de langage (LLM) afin d'aboutir à un modèle aux retours émotionnels variés et personnalisés pour correspondre au profil recherché par l'utilisateur.

Cet agent conversationnel émotionnellement personnalisé permettra *in fine* d'améliorer différents cas d'usages allant de l'aide médicale avec support émotionnel pour personnes en détresse ou en dysrégulation émotionnelle, à l'amélioration globale de l'interaction humain-machine par la personnalisation des retours émotionnels au travers de l'interaction avec leur recommandation.

2 Objectifs

Dans cette thèse, nous émettons l'hypothèse selon laquelle les modèles de langage peuvent être transformés en modèles d'agent conversationnels au comportement et retours émotionnels personnalisables. Ainsi, nous identifions trois objectifs principaux pour ce travail de thèse. Le premier objectif est le retour d'information émotionnelle, ou *emotional feedback* en anglais (Cassell and Thorisson, 1999), au travers d'un système de recommandation dédié. Pour cet objectif, nous nous concentrerons essentiellement sur la modalité textuelle afin de créer un assistant de retour d'information émotionnelle qui servira au second objectif qu'est la personnalisation. En effet, à l'aide du système de recommandation nous cherchons à personnaliser le modèle de langage par affinage et utilisation de personas (Pradhan and Lazar, 2021 ; Hao

and Kong, 2025) afin créer un LLM émotionnel les exploitant (Yang et al., 2024). Cet LLM sera ainsi personnalisable en continu par l'utilisateur humain, qui pourra ainsi façonner son propre agent empathique, ce qui va donc dans le sens inverse de travaux récents (Li et al., 2023).

Enfin, le dernier objectif de cette thèse est d'aller au delà d'un LLM textuel en y incluant une modalité supplémentaire : l'image. Ou, plus précisément, une représentation graphique du visage et de son retour d'information émotionnelle. Ainsi, nous associerons texte et visage virtuel en commençant par une représentation caricaturale des émotions avant d'y intégrer une représentation plus fine des indices visuels pour chaque émotion.

Bien entendu, afin de répondre à la question de recherche de savoir si un LLM émotionnel est personnalisable plus efficacement au travers d'indices uniquement textuels, ou d'indices faciaux, nous comparerons les processus d'affinage et de personnalisation par recommandation d'émotions dans la conversation au travers des deux modalités. De plus, cette thèse permettra de dépasser les processus d'entraînements opaques en redonnant, au moins partiellement, le contrôle de la personnalisation du modèle à l'utilisateur néophyte.

3 Méthodologie

Cette thèse sera réalisée de manière itérative. Les objectifs présentés précédemment peuvent être vus comme des étapes de réalisation qui se suivent et dépendent chacune de l'étape précédente. Nous aborderons ainsi dans un premier temps (1) la recommandation du retour d'information émotionnelle, avant de l'étoffer avec une (2) personnalisation émotionnelle du modèle par l'utilisateur. Puis nous finirons par la réutilisation des deux précédentes étapes pour (3) une mise en place des retours d'information multimodaux (texte et vision). Chaque étape possède ses propres enjeux méthodologiques que nous présentons ci-dessous.

Recommandation de retour d'information émotionnelle pour l'affinage du modèle par renforcement. Le premier enjeu de cette thèse est de permettre à l'agent conversationnel de fournir un retour d'information émotionnel. En effet, lors des échanges humain-machine, il est souvent nécessaire pour l'Agent Conversationnel Animé (ACA) de prédire la bonne émotion qui transparait au travers de la réponse de l'utilisateur (Axelsson et al., 2022). Cependant, une telle prédiction à des fins de détection d'émotions n'est pas suffisante dans notre cas : il est nécessaire de recommander des émotions en réponse à l'utilisateur pour lui permettre de garder la main sur l'évolution de l'agent empathique. Cet objectif sera ainsi conduit à l'aide de l'affinage (*finetuning*) de LLMs pré-entraînés. Surtout, ce travail se distingue par l'inclusion du retour utilisateur pour orienter l'affinage du modèle de langage à partir du retour d'information émotionnel de l'agent. Il conviendra d'utiliser des approches dites *human in the loop* (Slade et al., 2024) permettant, dans notre cas, à l'utilisateur de régénérer la dernière utterance donnée avec une émotion distincte. Il sera notamment nécessaire de mettre en place une interface utilisateur facilitant l'affinage interactif du modèle. Heureusement, de tels outils sont désormais légion¹ (Pei et al., 2022).

Personnalisation et profil utilisateur. Réutilisant l'affinage par *human-in-the-loop*, nous chercherons ici à attribuer dynamiquement une persona à l'utilisateur afin d'obtenir non seulement un modèle conversationnel issu de personas (Li et al., 2016) mais également une recommandation d'émotions et de retour d'information en fonction de l'utilisateur. Nous chercherons à vérifier si l'utilisation de personas permet également au modèle de respecter les données privées de l'utilisateur tout en lui conférant une meilleure explicabilité de la recommandation. Les personas pourront ainsi faire l'objet d'*adapters* (Houlsby et al., 2019) dédiés pour influencer le modèle de recommandation de retour d'information émotionnelle pré-affiné. Le tout afin d'obtenir un meilleur contrôle sur les réactions de l'agent conversationnel.

Modélisation visuelle des expressions des émotions de l'agent. Pour obtenir un agent conversationnel animé (ACA), il convient d'aborder le retour visuel fourni à l'utilisateur afin de montrer automatiquement

1. <https://www.gradio.app/docs>

les émotions ciblées par le modèle. Et ce, directement à partir d’une modélisation automatique des expressions du visage (Tisserand et al., 2020). Nous ajouterons donc ici la modalité visuelle au texte, en exploitant et améliorant les Vision Language Models (VLM) (Lu et al., 2019; Bordes et al., 2024) pour obtenir un modèle de bout à bout (*end-to-end*) permettant d’être à la fois personnalisation par interaction humaine et à partir de retours visuels. L’impact de retours multimodaux de la part des agents conversationnels a en effet montré une importance capitale dans la perception de ces agents par les humains (Elfleet and Chollet, 2024). La comparaison entre le VLM et le LLM permettra alors de répondre à notre hypothèse principale énoncée précédemment (Section 2).

4 Équipe d’encadrement

Davide Buscaldi – directeur de thèse. Davide Buscaldi est maître de conférences en informatique au laboratoire d’informatique de Paris Nord (LIPN) et tout nouvellement HDR depuis mars 2025. Il a déjà travaillé sur les aspects émotionnels dans ses recherches (Ghanem et al., 2020; García-Mendoza and Buscaldi, 2023) ainsi que sur la personnalisation et le profilage utilisateur (Palomino-Garibay et al., 2015). Enfin, par ses récents travaux sur les VLM (Dhouib et al., 2025), Davide Buscaldi apporte la multimodalité dans ce travail de thèse.

Gaël Guibon – co-encadrant. Gaël Guibon est maître de conférences en informatique au LIPN depuis septembre 2024. Sa recherche porte principalement sur les émotions et les conversations (Guibon et al., 2021b,a, 2023; Gendron and Guibon, 2024), tout en exploitant et étudiant l’usage de LLMs (Waffo Dzuyo et al., 2025; Ghebriout et al., 2025) pour les aspects conversationnels. Enfin, Gaël Guibon est notamment membre de l’association pour le traitement automatique des langues (ATALA) et de l’association sur les Affects, Compagnons Artificiels et Interactions (ACAI) et organisateur des groupes des ateliers ACAI et de la conférence majeure en TAL (ACL 2024, ACL 2026), dans laquelle ses travaux sur les émojis et émotions s’inscrivent également (Guibon et al., 2018).

Références

- Axelsson, A., Buschmeier, H., and Skantze, G. (2022). Modeling feedback in interaction with conversational agents—a review. *Frontiers in Computer Science*, 4 :744574.
- Bordes, F., Pang, R. Y., Ajay, A., Li, A. C., Bardes, A., Petryk, S., Mañas, O., Lin, Z., Mahmoud, A., Jayaraman, B., et al. (2024). An introduction to vision-language modeling. *arXiv preprint arXiv :2405.17247*.
- Brave, S. and Nass, C. (2007). Emotion in human-computer interaction. In *The human-computer interaction handbook*, pages 103–118. CRC Press.
- Cassell, J. and Thorisson, K. R. (1999). The power of a nod and a glance : Envelope vs. emotional feedback in animated conversational agents. *Applied Artificial Intelligence*, 13(4-5) :519–538.
- Dhouib, M., Buscaldi, D., Vanier, S., and Shabou, A. (2025). Pact : Pruning and clustering-based token reduction for faster visual language models. In *The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2025)*.
- Elfleet, M. and Chollet, M. (2024). Investigating the impact of multimodal feedback on user-perceived latency and immersion with llm-powered embodied conversational agents in virtual reality. In *Proceedings of the 24th ACM International Conference on Intelligent Virtual Agents, IVA ’24*, New York, NY, USA. Association for Computing Machinery.
- García-Mendoza, J.-L. and Buscaldi, D. (2023). Enriching with minority instances a corpus of sentiment analysis in spanish. In *IberLEF@ SEPLN*.

- Gendron, B. and Guibon, G. (2024). SEC : Context-aware metric learning for efficient emotion recognition in conversation. In De Clercq, O., Barriere, V., Barnes, J., Klinger, R., Sedoc, J., and Tafreshi, S., editors, *Proceedings of the 14th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis*, pages 11–22, Bangkok, Thailand. Association for Computational Linguistics.
- Ghanem, B., Buscaldi, D., and Rosso, P. (2020). Textrolls : Identifying trolls on twitter with textual and affective features. In *OHARS@ RecSys*, pages 4–22.
- Ghebruiout, M. I. E., Guibon, G., Lerner, I., and Vincent, E. (2025). Quartz : Approche abstractive non supervisée par question-réponse pour le résumé de dialogue orienté tâche. In *Actes de la 32ème Conférence sur le Traitement Automatique des Langues Naturelles, volume 1 : articles longs et prises de position*, Marseille, France. ATALA and ARIA.
- Guibon, G., Labeau, M., Flamein, H., Lefevre, L., and Clavel, C. (2021a). Few-shot emotion recognition in conversation with sequential prototypical networks. In Moens, M.-F., Huang, X., Specia, L., and Yih, S. W.-t., editors, *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 6858–6870, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.
- Guibon, G., Labeau, M., Flamein, H., Lefevre, L., and Clavel, C. (2021b). Meta-learning for classifying previously unseen data source into previously unseen emotional categories. In Lee, H.-Y., Mohtarami, M., Li, S.-W., Jin, D., Korpusik, M., Dong, S., Vu, N. T., and Hakkani-Tur, D., editors, *Proceedings of the 1st Workshop on Meta Learning and Its Applications to Natural Language Processing*, pages 76–89, Online. Association for Computational Linguistics.
- Guibon, G., Labeau, M., Lefevre, L., and Clavel, C. (2023). An adaptive layer to leverage both domain and task specific information from scarce data. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 7757–7765.
- Guibon, G., Ochs, M., and Bellot, P. (2018). Emoji recommendation in private instant messages. In *Proceedings of the 33rd annual acm symposium on applied computing*, pages 1821–1823.
- Hao, J. and Kong, F. (2025). Enhancing emotional support conversations : A framework for dynamic knowledge filtering and persona extraction. In Rambow, O., Wanner, L., Apidianaki, M., Al-Khalifa, H., Eugenio, B. D., and Schockaert, S., editors, *Proceedings of the 31st International Conference on Computational Linguistics*, pages 3193–3202, Abu Dhabi, UAE. Association for Computational Linguistics.
- Houlsby, N., Giurui, A., Jastrzebski, S., Morrone, B., De Laroussilhe, Q., Gesmundo, A., Attariyan, M., and Gelly, S. (2019). Parameter-efficient transfer learning for nlp. In *International conference on machine learning*, pages 2790–2799. PMLR.
- Kochan, J. (2013). Subjectivity and emotion in scientific research. *Studies in History and Philosophy of Science Part A*, 44(3) :354–362.
- Li, J., Galley, M., Brockett, C., Spithourakis, G. P., Gao, J., and Dolan, B. (2016). A persona-based neural conversation model. *arXiv preprint arXiv :1603.06155*.
- Li, Y., Inoue, K., Tian, L., Fu, C., Ishi, C. T., Ishiguro, H., Kawahara, T., and Lai, C. (2023). I know your feelings before you do : Predicting future affective reactions in human-computer dialogue. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–7.
- Lu, J., Batra, D., Parikh, D., and Lee, S. (2019). Vilbert : Pretraining task-agnostic visiolinguistic representations for vision-and-language tasks. *Advances in neural information processing systems*, 32.

- Palomino-Garibay, A., Camacho-Gonzalez, A. T., Fierro-Villaneda, R. A., Hernandez-Farias, I., Buscaldi, D., Meza-Ruiz, I. V., et al. (2015). A random forest approach for authorship profiling. In *Proceedings of CLEF*.
- Pei, J., Ananthasubramaniam, A., Wang, X., Zhou, N., Dedeloudis, A., Sargent, J., and Jurgens, D. (2022). Potato : The portable text annotation tool. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing : System Demonstrations*.
- Picard, R. W. (2000). *Affective computing*. MIT press.
- Pradhan, A. and Lazar, A. (2021). Hey google, do you have a personality ? designing personality and personas for conversational agents. In *Proceedings of the 3rd Conference on Conversational User Interfaces*, pages 1–4.
- Slade, P., Atkeson, C., Donelan, J. M., Houdijk, H., Ingraham, K. A., Kim, M., Kong, K., Poggensee, K. L., Riener, R., Steinert, M., et al. (2024). On human-in-the-loop optimization of human–robot interaction. *Nature*, 633(8031) :779–788.
- Tisserand, Y., Aylett, R., Mortillaro, M., and Rudrauf, D. (2020). Real-time simulation of virtual humans’ emotional facial expressions, harnessing autonomic physiological and musculoskeletal control. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents, IVA ’20*, New York, NY, USA. Association for Computing Machinery.
- Waffo Dzuyo, G. S., Guibon, G., Cerisara, C., and Belmar-Letelier, L. (2025). Linking industry sectors and financial statements : A hybrid approach for company classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 16444–16452.
- Yang, B., Liu, D., Xiao, C., Zhao, K., Tang, C., Li, C., Yuan, L., Yang, G., Huang, L., and Lin, C. (2024). Crafting customisable characters with llms : Introducing simschat, a persona-driven role-playing agent framework. *arXiv preprint arXiv :2406.17962*.