

Manying Zhang

Agents conversationnels personnalisés et multi-modaux dans le domaine de la mode

Directeur de thèse :

Mathieu Valette, Damien Nouvel



Date de soutenance :

Le 16 janvier 2026

Résumé

Dans le cadre de cette thèse, nous étudions la conception et le développement d'un agent conversationnel personnalisé et multimodal pour le commerce électronique dans le domaine de la mode. Notre objectif principal est de construire un système capable de comprendre les requêtes complexes des utilisateurs, de recommander des produits pertinents, et de répondre de manière naturelle tout en intégrant des données multimodales (texte, image) et des caractéristiques personnalisées des utilisateurs.

La méthodologie adoptée repose sur trois piliers fondamentaux. Tout d'abord, une stratégie rigoureuse de collecte et génération de données a été mise en place, combinant l'exploitation de corpus existants, la génération automatique via des simulateurs d'utilisateurs et de systèmes, ainsi que des données collectées en ligne. À partir de ces sources, un jeu de données conversationnel dans le domaine de la mode a été constitué, accompagné d'une base de connaissances, comportant des informations produits détaillées, des profils utilisateurs, ainsi que des interactions complexes.

Ensuite, nous avons développé un système modulaire reposant sur plusieurs types de modèles de fondation : des classifieurs, des modèles d'embeddings et des grands modèles de langage (LLM). Ces modèles ont été entraînés ou adaptés individuellement, puis orchestrés dans un pipeline de dialogue structuré, intégrant des modules spécialisés pour la compréhension, la récupération d'information, la génération de réponses, et la personnalisation. Différentes architectures ont été comparées, incluant des approches séquentielles et des modèles de bout-en-bout.

Enfin, nous avons exploré un mécanisme original d'auto-alignement itératif, combinant deux simulateurs : un simulateur d'utilisateur, qui modélise les préférences et les comportements d'achat, et un simulateur de vendeur, qui génère des descriptions de produits à partir de la perspective utilisateur. En confrontant ces deux pôles, le système apprend à générer un langage persuasif pour les produits décrits, et à générer des descriptions en fonction de l'utilité perçue par l'utilisateur simulé.



Institut national
des langues
et civilisations orientales

Les expériences menées montrent que ce système permet une interaction plus personnalisée et pertinente, tout en mettant en lumière les défis méthodologiques liés à la fois à la mise en œuvre d'un système de dialogue, à sa personnalisation et à l'intégration de multimodalité. Ce travail ouvre des perspectives sur le développement d'agents conversationnels dans le cadre d'applications commerciales réelles.