
Interpréter une variable latente avec un modèle de langage, une application du package NaileR à la compréhension des systèmes alimentaires durables

Nel Hervé*¹, Sylvie Thoby*¹, and Sébastien Lê*^{2,3}

¹L’Institut Agro Rennes-Angers – L’Institut Agro Agrocampus Ouest – 65, rue de Saint-Brieuc, 35000 Rennes Cedex, France

²Institut de recherche mathématique de Rennes (IRMAR) – CNRS – France

³L’Institut Agro Rennes-Angers – L’Institut Agro Agrocampus Ouest – 65, rue de Saint-Brieuc, 35000 Rennes Cedex, France

Résumé

La Q-méthode ou Q-méthodologie est une méthode de recherche utilisée en sciences humaines et sociales pour étudier des concepts complexes à travers les points de vue subjectifs d’individus interrogés. Pour ce faire, on identifie un ensemble de stimuli, généralement des assertions, que l’on propose aux interrogés de trier dans des groupes définis a priori par leur sens et par le nombre d’assertions qu’ils peuvent contenir. Dans notre recherche, nous nous intéressons aux systèmes alimentaires durables et en particulier à la perception que l’on peut en avoir, suivant que l’on privilégie la dimension économique, sociale, environnementale...

Pour comprendre cette variabilité de perception, nous avons soumis un ensemble de 45 assertions, structurées selon 5 dimensions, à plus de 500 personnes : un système alimentaire qui répondrait à mes valeurs et besoins, sans compromettre la capacité des générations futures à répondre aux leurs, vendrait beaucoup plus cher les produits qui ne sont pas de saison, imposerait une redistribution équitable des bénéfices tout au long de la chaîne alimentaire, limiterait la viande à deux jours par semaine dans la restauration collective,...

Les données recueillies ont permis d’obtenir une typologie d’interrogés, telle que deux personnes interrogées appartiennent à la même classe si elles partagent la même perception d’un système alimentaire durable. L’interprétation des classes issues de la typologie a été faite “manuellement” d’une part, à l’aide du package NaileR d’autre part. NaileR est un package R conçu pour interpréter des variables latentes continues ou catégorielles : typiquement, des dimensions issues d’une analyse multivariée exploratoire, ou une variable de classe issue d’un algorithme de classification non supervisée. Dans cette recherche, nous montrerons comment tester la similarité entre l’interprétation “manuelle” et les interprétations fournies par un modèle de langage tel que Llama3. Nous proposerons également une représentation graphique de ces différentes interprétations.

Mots-Clés: Llama3, Q méthode, système alimentaire, similarité textuelle

*Intervenant