Research Article

New Approaches for Food Marketing Strategies and Sectoral Policies: Text Mining and NLP-Based Sentiment Analysis

Yusuf ÇAKMAKÇI*[1], Nuray HELVACI[2], Harun HURMA[3], İdris AKKUZU[4]

[1,2,3]Department of Agricultural Economics, Faculty of Agriculture, Tekirdağ Namık Kemal University, Tekirdağ, Türkiye
[2,4]Department of Political Science and Public Administration, Faculty of Economics and Administrative Sciences, Tekirdağ Namık Kemal University, Tekirdağ, Türkiye

[1]https://orcid.org/0000-0002-5136-9102, [2]https://orcid.org/0009-0003-8602-5099, [3]https://orcid.org/0000-0003-1845-3940
[4]https://orcid.org/0000-0003-1069-9719

*Corresponding author e-mail: cak76yusuf@gmail.com

Abstract: This study investigates the application of Natural Language Processing (NLP) and text mining techniques to enhance the understanding of consumer behavior and sectoral perspectives within the food marketing and horse racing industries. Two primary datasets were utilized: a face-to-face survey conducted with 171 consumers in Tekirdağ, Türkiye, and in-depth interviews with 20 employees from the Turkish horse racing sector. Survey responses were analyzed using R, incorporating text mining methods such as word frequency analysis, bigram identification, chi-square testing, and network analysis. The findings revealed statistically significant associations between purchased foods and specific demographic variables. Notably, household size was significantly associated with cheese consumption ($\chi^2(1) = 6.453$, $p = 0.011$), and gender was significantly related to vegetable consumption ($\chi^2(1) = 4.168$, $p = 0.041$). Additionally, borderline associations were identified between gender and fruit ($p = 0.061$) and egg ($p = 0.080$) consumption, as well as between the number of household workers and yoghurt consumption ($p = 0.054$). Network analysis highlighted the central role of items such as vegetables, fruit, milk, and cheese across various labeling categories, including "organic," "natural," and "cooperative." Interview data were processed in Python using sentiment analysis and clustering techniques. Two primary sentiment-based clusters emerged: one reflecting positive perceptions related to horse care and professional identity, and another indicating dissatisfaction with social life and work-life balance. Overall, the study emphasizes the importance of natural language processing (NLP) and text mining in producing reliable information to influence marketing strategies and policy making in agricultural economics and social science fields.