

ing terms for web pages. *Computer Networks* 33, 1-6 (2000), 247–255.

- [16] VENABLES, W. N., RIPLEY, B. D., VENABLES, W., AND RIPLEY, B. Tree-based methods. *Modern applied statistics with S-Plus* (1999), 303–327.
- [17] WANG, L., YANG, N., HUANG, X., YANG, L., MAJUMDER, R., AND WEI, F. Improving text embeddings with large language models, 2024.
- [18] XIA, P., ZHANG, L., AND LI, F. Learning similarity with cosine similarity ensemble. *Information sciences* 307 (2015), 39–52.

DRAFT