

that even though a great amount of work on natural language processing is available in literature surveys (one may refer to [15, 32, 63, 98, 133, 151] focusing on one domain such as usage of deep-learning techniques in NLP, techniques used for email spam filtering, medication safety, management research, intrusion detection, and Gujarati language etc.), still there is not much work on regional languages, which can be the focus of future research.

**Acknowledgements** Authors would like to express the gratitude to Research Mentors from CL Educate: Accendere Knowledge Management Services Pvt. Ltd. for their comments on earlier versions of the manuscript. Although any errors are our own and should not tarnish the reputations of these esteemed persons. We would also like to appreciate the Editor, Associate Editor, and anonymous referees for their constructive suggestions that led to many improvements on an earlier version of this manuscript.

## Declarations

**Conflict of interest** The first draft of this paper was written under the supervision of Dr. Kiran Khatter and Dr. Sukhdev Singh, associated with CL- Educate: Accendere Knowledge Management Services Pvt. Ltd. and deputed at the Manav Rachna International University. The draft is also available on [arxiv.org](https://arxiv.org/abs/1708.05148) at <https://arxiv.org/abs/1708.05148>

## References

1. Ahonen H, Heinonen O, Klemettinen M, Verkamo AI (1998) Applying data mining techniques for descriptive phrase extraction in digital document collections. In research and technology advances in digital libraries, 1998. ADL 98. Proceedings. IEEE international forum on (pp. 2-11). IEEE
2. Alshawi H (1992) The core language engine. MIT press
3. Alshemali B, Kalita J (2020) Improving the reliability of deep neural networks in NLP: A review. *Knowl-Based Syst* 191:105210
4. Andreev ND (1967) The intermediary language as the focal point of machine translation. In: Booth AD (ed) Machine translation. North Holland Publishing Company, Amsterdam, pp 3–27
5. Androutsopoulos I, Paliouras G, Karkaletsis V, Sakkis G, Spyropoulos CD, Stamatopoulos P (2000) Learning to filter spam e-mail: A comparison of a naive bayesian and a memory-based approach. *arXiv preprint cs/0009009*
6. Baclic O, Tunis M, Young K, Doan C, Swerdfeger H, Schonfeld J (2020) Artificial intelligence in public health: challenges and opportunities for public health made possible by advances in natural language processing. *Can Commun Dis Rep* 46(6):161
7. Bahdanau D, Cho K, Bengio Y (2015) Neural machine translation by jointly learning to align and translate. In *ICLR 2015*
8. Bangalore S, Rambow O, Whittaker S (2000) Evaluation metrics for generation. In *proceedings of the first international conference on natural language generation-volume 14* (pp. 1-8). Assoc Comput Linguist
9. Baud RH, Rassinoux AM, Scherrer JR (1991) Knowledge representation of discharge summaries. In *AIME 91* (pp. 173–182). Springer, Berlin Heidelberg
10. Baud RH, Rassinoux AM, Scherrer JR (1992) Natural language processing and semantical representation of medical texts. *Methods Inf Med* 31(2):117–125
11. Baud RH, Alpay L, Lovis C (1994) Let's meet the users with natural language understanding. *Knowledge and Decisions in Health Telematics: The Next Decade* 12:103
12. Bengio Y, Ducharme R, Vincent P (2001) A neural probabilistic language model. *Proceedings of NIPS*
13. Benson E, Haghighi A, Barzilay R (2011) Event discovery in social media feeds. In *proceedings of the 49th annual meeting of the Association for Computational Linguistics: human language technologies-volume 1* (pp. 389-398). Assoc Comput Linguist
14. Berger AL, Della Pietra SA, Della Pietra VJ (1996) A maximum entropy approach to natural language processing. *Computational Linguistics* 22(1):39–71
15. Blanzieri E, Bryl A (2008) A survey of learning-based techniques of email spam filtering. *Artif Intell Rev* 29(1):63–92