**REFERENCES**

[1] P. K. Roy, J. P. Singh, and S. Banerjee, ``Deep learning to \_lter SMS spam,''

*Future Gener. Comput. Syst.*, vol. 102, pp. 524\_533, Jan. 2020.

[2] G. Jain, M. Sharma, and B. Agarwal, ``Optimizing semantic LSTM for

spam detection,'' *Int. J. Inf. Technol.*, vol. 11, no. 2, pp. 239\_250, Jun. 2019.

[3] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez,

L. Kaiser, and I. Polosukhin, ``Attention is all you need,'' in *Proc. Adv.*

*Neural Inf. Process. Syst.*, 2017, pp. 5999\_6009.

[4] T. B. Brown *et al.*, ``Language models are few-shot learners,'' 2020,

*arXiv:2005.14165*. [Online]. Available: http://arxiv.org/abs/2005.14165

[5] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, ``BERT: Pre-training

of deep bidirectional transformers for language understanding,'' in *Proc.*

*Conf. North Amer. Chapter Assoc. Comput. Linguistics, Hum. Lang. Tech-*

*nol.*, vol. 1, Jun. 2019, pp. 4171\_4186.

[6] G. Sonowal and K. S. Kuppusamy, ``SmiDCA: An anti-Smishing

model with machine learning approach,'' *Comput. J.*, vol. 61, no. 8,

pp. 1143\_1157, Aug. 2018.

[7] J. W. Joo, S. Y. Moon, S. Singh, and J. H. Park, ``S-detector: An enhanced

security model for detecting Smishing attack for mobile computing,''

*Telecommun. Syst.*, vol. 66, no. 1, pp. 29\_38, Sep. 2017.

[8] S. Mishra and D. Soni, ``Smishing detector: A security model to detect

Smishing through SMS content analysis and URL behavior analysis,''

*Future Gener. Comput. Syst.*, vol. 108, pp. 803\_815, Jul. 2020.

[9] C. Li, L. Hou, B. Y. Sharma, H. Li, C. Chen, Y. Li, X. Zhao, H. Huang,

Z. Cai, and H. Chen, ``Developing a new intelligent system for the diagnosis

of tuberculous pleural effusion,'' *Comput. Methods Programs Biomed.*,

vol. 153, pp. 211\_225, Jan. 2018.

[10] T. K. Ho, ``Random decision forests,'' in *Proc. Int. Conf. Document Anal.*

*Recognit. (ICDAR)*, vol. 1, 1995, pp. 278\_282.

[11] C. Cortes and V. Vapnik, ``Support-vector networks,'' *Mach. Learn.*,

vol. 20, no. 3, pp. 273\_297, 1995.

[12] M. Gupta, A. Bakliwal, S. Agarwal, and P. Mehndiratta, ``A comparative

study of spam SMS detection using machine learning classi\_ers,'' in *Proc.*

*11th Int. Conf. Contemp. Comput. (IC3)*, Aug. 2018, pp. 1\_7.

[13] T. A. Almeida, J. M. G. Hidalgo, and A. Yamakami, ``Contributions to the

study of SMS spam \_ltering: New collection and results,'' in *Proc. 11th*

*ACM Symp. Document Eng.*, Sep. 2011, pp. 259\_262.

[14] A. K. Jain and B. B. Gupta, ``Rule-based framework for detection of Smishing

messages in mobile environment,'' *Procedia Comput. Sci.*, vol. 125,

pp. 617\_623, 2018.

[15] W. W. Cohen, ``Fast effective rule induction,'' in *Machine Learning Pro-*

*ceedings*, 1995, pp. 115\_123.

[16] J. Cendrowska, ``PRISM: An algorithm for inducing modular rules,'' *Int.*

*J. Man-Machine Stud.*, vol. 27, no. 4, pp. 349\_370, Oct. 1987.

[17] J. H. Friedman, ``Greedy function approximation: A gradient boosting

machine,'' *Ann. Statist.*, vol. 29, no. 5, pp. 1189\_1232, Oct. 2001.

[18] L. Bottou, ``Large-scale machine learning with stochastic gradient

descent,'' in *Proc. COMPSTAT*. Physica-Verlag, 2010, pp. 177\_186.

[19] T. Mikolov, K. Chen, G. S. Corrado, and J. Dean, ``Ef\_cient estimation of

word representations in vector space,'' in *Proc. Int. Conf. Learn. Repre-*

*sent.*, 2013.

[20] G. A. Miller, ``WordNet: A lexical database for English,'' *Commun. ACM*,

vol. 38, no. 11, pp. 39\_41, 1995.

[21] H. Liu and P. Singh, ``ConceptNet\_A practical commonsense reasoning

tool-kit,'' *BT Technol. J.*, vol. 22, no. 4, pp. 211\_226, Oct. 2004.

[22] A. Ghourabi, M. A. Mahmood, and Q. M. Alzubi, ``A hybrid CNN-LSTM

model for SMS spam detection in arabic and English messages,'' *Future*

*Internet*, vol. 12, no. 9, p. 156, Sep. 2020.

[23] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, ``Learning representations

by back-propagating errors,'' *Nature*, vol. 323, no. 6088,

pp. 533\_536, Oct. 1986.

[24] Y. Bengio, P. Simard, and P. Frasconi, ``Learning long-term dependencies

with gradient descent is dif\_cult,'' *IEEE Trans. Neural Netw.*, vol. 5, no. 2,

pp. 157\_166, Mar. 1994.

[25] R. Pascanu, T. Mikolov, and Y. Bengio, ``On the dif\_culty of training

recurrent neural networks,'' in *Proc. 30th Int. Conf. Mach. Learn. (ICML)*,

2013, pp. 2347\_2355.