

REFERENCES

- [1] P. S. Kochhar, D. Lo, J. Lawall, and N. Nagappan, "Code coverage and postrelease defects: A large-scale study on open source projects," *IEEE Transactions on Reliability*, vol. 66, no. 4, pp. 1213–1228, 2017.
- [2] T. T. Chekam, M. Papadakis, Y. Le Traon, and M. Harman, "An empirical study on mutation, statement and branch coverage fault revelation that avoids the unreliable clean program assumption," *2017 IEEE/ACM 39th International Conference on Software Engineering (ICSE)*, 2017.
- [3] V. H. Durelli, R. S. Durelli, S. S. Borges, A. T. Endo, M. M. Eler, D. R. Dias, and M. P. Guimaraes, "Machine learning applied to software testing: A systematic mapping study," *IEEE Transactions on Reliability*, vol. 68, no. 3, pp. 1189–1212, 2019.
- [4] K. Sneha and G. M. Malle, "Research on software testing techniques and software automation testing tools," *2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*, 2017.
- [5] M. M. Almasi, H. Hemmati, G. Fraser, A. Arcuri, and J. Benefelds, "An industrial evaluation of UNIT TEST GENERATION: Finding real faults in a financial application," *2017 IEEE/ACM 39th International Conference on Software Engineering: Software Engineering in Practice Track (ICSE-SEIP)*, 2017.
- [6] G. Grano, T. V. Titov, S. Panichella, and H. C. Gall, "How high will it be? using machine learning models to predict branch coverage in automated testing," *2018 IEEE Workshop on Machine Learning Techniques for Software Quality Evaluation (MaLTesQuE)*, 2018.
- [7] A. Panichella, J. Campos, and G. Fraser, "EvoSuite at the SBST 2020 Tool Competition," *Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops*, 2020.
- [8] A. K. Uysal and S. Gunal, "The impact of preprocessing on text classification," *Information Processing & Management*, vol. 50, no. 1, pp. 104–112, 2014.
- [9] D. Hussien Maulud, S. R. Zeebaree, K. Jacksi, M. A. Mohammed Sadeeq, and K. Hussein Sharif, "State of art for Semantic Analysis of Natural Language Processing," *Qubahan Academic Journal*, vol. 1, no. 2, 2021.
- [10] E. Cueva, G. Ee, A. Iyer, A. Pereira, A. Roseman, and D. Martinez, "Detecting fake news on Twitter using Machine Learning Models," *2020 IEEE MIT Undergraduate Research Technology Conference (URTC)*, 2020.
- [11] J. Reyes, D. Ramirez, and J. Paciello, "Automatic classification of source code archives by programming language: A deep learning approach," *2016 International Conference on Computational Science and Computational Intelligence (CSCI)*, 2016.
- [12] S. Gilda, "Source code classification using Neural Networks," *2017 14th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, 2017.
- [13] W. Zheng, J. Gao, X. Wu, F. Liu, Y. Xun, G. Liu, and X. Chen, "The impact factors on the performance of machine learning-based Vulnerability Detection: A Comparative Study," *Journal of Systems and Software*, vol. 168, p. 110659, 2020.
- [14] J. Lilleberg, Y. Zhu, and Y. Zhang, "Support Vector Machines and word2vec for text classification with Semantic Features," *2015 IEEE 14th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC)*, 2015.
- [15] Y. Goldberg and O. Levy, "word2vec Explained: deriving Mikolov et al.'s negative-sampling word-embedding method", *arXiv preprint arXiv:1402.3722*, 2014.
- [16] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, "Scikit-learn: Machine learning in Python", *the Journal of machine Learning research*, vol. 12, pp. 2825–2830, 2011.
- [17] S. Ugurel, R. Krovetz, and C. L. Giles, "What's the code? automatic classification of source code archives," *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '02*, 2002.
- [18] R. Puri, D. S. Kung, G. Janssen, W. Zhang, G. Domeniconi, V. Zolotov, J. Dolby, J. Chen, M. Choudhury, L. Decker *et al.*, "Project codenet: A large-scale ai for code dataset for learning a diversity of coding tasks", *arXiv preprint arXiv:2105.12655*, vol. 1035, 2021.