

DAFTAR PUSTAKA

- [1] “State of Agile Report,” Digital AI. Accessed: Nov. 14, 2024. [Online]. Available: <https://info.digital.ai/rs/981-LQX-968/images/AR-SA-2022-16th-Annual-State-Of-Agile-Report.pdf>
- [2] SCRUMstudy™, *A guide to the Scrum Body of knowledge (SBOK Guide)*, 2013 edition. Phoenix, Arizona: SCRUMstudy, A brand of VMEdU, Inc, 2013.
- [3] M. Choetkiertikul, H. K. Dam, T. Tran, T. Pham, A. Ghose, and T. Menzies, “A Deep Learning Model for Estimating Story Points,” *IEEE Trans. Softw. Eng.*, vol. 45, no. 7, pp. 637–656, July 2019, doi: 10.1109/TSE.2018.2792473.
- [4] Ritu and Y. Garg, “Comparative Analysis of Machine Learning Techniques in Effort Estimation,” in *2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON)*, Faridabad, India: IEEE, May 2022, pp. 401–405. doi: 10.1109/COM-IT-CON54601.2022.9850592.
- [5] N. Govil and A. Sharma, “Estimation of cost and development effort in Scrum-based software projects considering dimensional success factors,” *Adv. Eng. Softw.*, vol. 172, p. 103209, Oct. 2022, doi: 10.1016/j.advengsoft.2022.103209.
- [6] C. Ratke, H. H. Hoffmann, T. Gaspar, and P. E. Floriani, “Effort Estimation using Bayesian Networks for Agile Development,” in *2019 2nd International Conference on Computer Applications & Information Security (ICCAIS)*, Riyadh, Saudi Arabia: IEEE, May 2019, pp. 1–4. doi: 10.1109/CAIS.2019.8769455.
- [7] R. K. Mallidi and M. Sharma, “Study on Agile Story Point Estimation Techniques and Challenges,” *Int. J. Comput. Appl.*, vol. 174, no. 13, pp. 9–14, Jan. 2021, doi: 10.5120/ijca2021921014.
- [8] T. Catak, P. O. Durdu, and S. Ilhan Omurca, “Enhancing Agile Effort Estimation: An NLP Approach for Software Requirements Analysis,” in *2024 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, Istanbul, Turkiye: IEEE, May 2024, pp. 1–8. doi: 10.1109/HORA61326.2024.10550870.
- [9] M. Usman, E. Mendes, F. Weidt, and R. Britto, “Effort estimation in agile software development: a systematic literature review,” in *Proceedings of the 10th International Conference on Predictive Models in Software Engineering*, Turin Italy: ACM, Sept. 2014, pp. 82–91. doi: 10.1145/2639490.2639503.
- [10] Y. Ramchurreetoo and V. Hurbungs, “A multiclass classification model to estimate Agile user stories,” in *2022 3rd International Conference on Next Generation Computing Applications (NextComp)*, Flic-en-Flac, Mauritius: IEEE, Oct. 2022, pp. 1–5. doi: 10.1109/NextComp55567.2022.9932190.
- [11] A. Zakrani, A. Najm, and A. Marzak, “Support Vector Regression Based on Grid-Search Method for Agile Software Effort Prediction,” in *2018 IEEE 5th International Congress on Information Science and Technology (CiSt)*, Marrakech: IEEE, Oct. 2018, pp. 1–6. doi: 10.1109/CIST.2018.8596370.
- [12] Z. Sakhravi, A. Sellami, and N. Bouassida, “Support vector regression for enhancement effort prediction of Scrum projects from COSMIC functional size,” *Innov. Syst. Softw. Eng.*, vol. 18, no. 1, pp. 137–153, Mar. 2022, doi: 10.1007/s11334-021-00420-8.
- [13] E. R. Sanchez, H. C. Maceda, and E. V. Santacruz, “Software Effort Estimation for Agile Software Development Using a Strategy Based on k-Nearest Neighbors Algorithm,” in *2022 IEEE Mexican International Conference on Computer Science*

- (ENC), Xalapa, Veracruz, Mexico: IEEE, Aug. 2022, pp. 1–6. doi: 10.1109/ENC56672.2022.9882947.
- [14] V. Tawosi, A. Al-Subaih, and F. Sarro, “Investigating the Effectiveness of Clustering for Story Point Estimation,” in *2022 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER)*, Honolulu, HI, USA: IEEE, Mar. 2022, pp. 827–838. doi: 10.1109/SANER53432.2022.00101.
- [15] Ch. Prasada Rao, P. Siva Kumar, S. Rama Sree, and J. Devi, “An Agile Effort Estimation Based on Story Points Using Machine Learning Techniques,” in *Proceedings of the Second International Conference on Computational Intelligence and Informatics*, vol. 712, V. Bhateja, J. M. R. S. Tavares, B. P. Rani, V. K. Prasad, and K. S. Raju, Eds., in *Advances in Intelligent Systems and Computing*, vol. 712., Singapore: Springer Singapore, 2018, pp. 209–219. doi: 10.1007/978-981-10-8228-3_20.
- [16] M. Hamid, F. Zeshan, and A. Ahmad, “Fuzzy Logic-based Expert System for Effort Estimation in Scrum Projects,” in *2021 International Conference on Decision Aid Sciences and Application (DASA)*, Sakheer, Bahrain: IEEE, Dec. 2021, pp. 761–765. doi: 10.1109/DASA53625.2021.9682239.
- [17] P. Hansen and H. Timinger, “Concept of a Fuzzy Expert System for Story Point Estimations in Agile Projects,” in *2022 IEEE 28th International Conference on Engineering, Technology and Innovation (ICE/ITMC) & 31st International Association For Management of Technology (IAMOT) Joint Conference*, Nancy, France: IEEE, June 2022, pp. 1–9. doi: 10.1109/ICE/ITMC-IAMOT55089.2022.10033271.
- [18] B. Marapelli, A. Carie, and S. M. N. Islam, “RNN-CNN MODEL:A Bi-directional Long Short-Term Memory Deep Learning Network For Story Point Estimation,” in *2020 5th International Conference on Innovative Technologies in Intelligent Systems and Industrial Applications (CITISIA)*, Sydney, Australia: IEEE, Nov. 2020, pp. 1–7. doi: 10.1109/CITISIA50690.2020.9371770.
- [19] M. Fu and C. Tantithamthavorn, “GPT2SP: A Transformer-Based Agile Story Point Estimation Approach,” *IEEE Trans. Softw. Eng.*, vol. 49, no. 2, pp. 611–625, Feb. 2023, doi: 10.1109/TSE.2022.3158252.
- [20] E. M. De Bortoli Fávero, D. Casanova, and A. R. Pimentel, “SE 3 M: A model for software effort estimation using pre-trained embedding models,” *Inf. Softw. Technol.*, vol. 147, p. 106886, July 2022, doi: 10.1016/j.infsof.2022.106886.
- [21] M. Alhamed and T. Storer, “Evaluation of Context-Aware Language Models and Experts for Effort Estimation of Software Maintenance Issues,” in *2022 IEEE International Conference on Software Maintenance and Evolution (ICSME)*, Limassol, Cyprus: IEEE, Oct. 2022, pp. 129–138. doi: 10.1109/ICSME55016.2022.00020.
- [22] V. Tawosi, A. Al-Subaih, R. Moussa, and F. Sarro, “A versatile dataset of agile open source software projects,” in *Proceedings of the 19th International Conference on Mining Software Repositories*, Pittsburgh Pennsylvania: ACM, May 2022, pp. 707–711. doi: 10.1145/3524842.3528029.
- [23] Y. Wu *et al.*, “Google’s Neural Machine Translation System: Bridging the Gap between Human and Machine Translation,” Oct. 08, 2016, *arXiv*: arXiv:1609.08144. doi: 10.48550/arXiv.1609.08144.
- [24] E. Sharkey and P. Treleaven, “Comparison of BERT vs GPT”.
- [25] J. Zhang, Y. Huang, S. Liu, Y. Gao, and X. Hu, “Do BERT-Like Bidirectional Models Still Perform Better on Text Classification in the Era of LLMs?,” 2025, *arXiv*. doi:

- 10.48550/ARXIV.2505.18215.
- [26] S. Porru, A. Murgia, S. Demeyer, M. Marchesi, and R. Tonelli, “Estimating Story Points from Issue Reports,” in *Proceedings of the The 12th International Conference on Predictive Models and Data Analytics in Software Engineering*, Ciudad Real Spain: ACM, Sept. 2016, pp. 1–10. doi: 10.1145/2972958.2972959.
- [27] M. Abadeer and M. Sabetzadeh, “Machine Learning-based Estimation of Story Points in Agile Development: Industrial Experience and Lessons Learned,” in *2021 IEEE 29th International Requirements Engineering Conference Workshops (REW)*, Notre Dame, IN, USA: IEEE, Sept. 2021, pp. 106–115. doi: 10.1109/REW53955.2021.00022.
- [28] H. Phan and A. Jannesari, “Story point level classification by text level graph neural network,” in *Proceedings of the 1st International Workshop on Natural Language-based Software Engineering*, Pittsburgh Pennsylvania: ACM, May 2022, pp. 75–78. doi: 10.1145/3528588.3528654.
- [29] H. Phan and A. Jannesari, “Heterogeneous Graph Neural Networks for Software Effort Estimation,” in *Proceedings of the 16th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement*, Helsinki Finland: ACM, Sept. 2022, pp. 103–113. doi: 10.1145/3544902.3546248.
- [30] B. Permana, R. Ferdiana, and A. Pratama, “Large Language Model Employment for Story Point Estimation Problems in AGILE Development,” in *2024 International Conference on Electrical Engineering and Computer Science (ICECOS)*, Palembang, Indonesia: IEEE, Sept. 2024, pp. 391–398. doi: 10.1109/ICECOS63900.2024.10791206.
- [31] A. Carrera-Rivera, W. Ochoa, F. Larrinaga, and G. Lasa, “How-to conduct a systematic literature review: A quick guide for computer science research,” *MethodsX*, vol. 9, p. 101895, 2022, doi: 10.1016/j.mex.2022.101895.
- [32] V. Tawosi, R. Moussa, and F. Sarro, “Agile Effort Estimation: Have We Solved the Problem Yet? Insights From a Replication Study,” *IEEE Trans. Softw. Eng.*, vol. 49, no. 4, pp. 2677–2697, Apr. 2023, doi: 10.1109/TSE.2022.3228739.
- [33] S. K. Tipu and S. Zia, “An Effort Estimation Model for Agile Software Development,” vol. 2, no. 1, 2012.
- [34] M. Alhamed and T. Storer, “JOSSE: A Software Development Effort Dataset Annotated with Expert Estimates.” Zenodo, Aug. 25, 2022. doi: 10.5281/ZENODO.7022735.
- [35] “ExampleApp Functional Specification | PDF | Use Case | Software Testing,” Scribd. Accessed: Dec. 15, 2024. [Online]. Available: <https://www.scribd.com/document/418340876/ExampleApp-Functional-Specification-xls>
- [36] M. Allamanis, M. Brockschmidt, and M. Khademi, “Learning to Represent Programs with Graphs,” May 04, 2018, *arXiv*: arXiv:1711.00740. doi: 10.48550/arXiv.1711.00740.
- [37] L. Huang, D. Ma, S. Li, X. Zhang, and H. WANG, “Text Level Graph Neural Network for Text Classification,” Oct. 08, 2019, *arXiv*: arXiv:1910.02356. doi: 10.48550/arXiv.1910.02356.
- [38] R. Sennrich, B. Haddow, and A. Birch, “Neural Machine Translation of Rare Words with Subword Units,” June 10, 2016, *arXiv*: arXiv:1508.07909. doi: 10.48550/arXiv.1508.07909.
- [39] J. Shore and S. Warden, *The art of agile development*, 1. ed. in Theory in practice. Beijing Köln: O’Reilly, 2008.

- [40] K. Schwaber and J. Sutherland, “The Scrum Guide. The Definitive Guide to Scrum: The Rules of the Game.” Accessed: Dec. 20, 2024. [Online]. Available: <https://scrumguides.org/docs/scrumguide/v2020/2020-Scrum-Guide-US.pdf>
- [41] M. Cohn, *User stories applied: for agile software development*, 18. print. in Addison-Wesley signature series. Boston, Mass.: Addison-Wesley, 2013.
- [42] G. Lucassen, F. Dalpiaz, J. M. E. M. Van Der Werf, and S. Brinkkemper, “Improving agile requirements: the Quality User Story framework and tool,” *Requir. Eng.*, vol. 21, no. 3, pp. 383–403, Sept. 2016, doi: 10.1007/s00766-016-0250-x.
- [43] M. Cohn, *Agile Estimating and Planning*. in Robert C. Martin. Pearson Education, 2005.
- [44] Dr. M. Singla, Ed., *Advances in Computer Science*, 1st ed. AkiNik Publications, 2020. doi: 10.22271/ed.book.784.
- [45] Farhanaaz and V. Sanju, “An exploration on lexical analysis,” in *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, Chennai, India: IEEE, Mar. 2016, pp. 253–258. doi: 10.1109/ICEEOT.2016.7755127.
- [46] N. Indurkha, Ed., *Handbook of natural language processing*. in 4190 Chapman & Hall/CRC machine learning & pattern recognition series. Boca Raton, FL: CRC Pr, 2010.
- [47] M. V. Redd, “Semantical and Syntactical Analysis of NLP,” vol. 5, 2014.
- [48] C. S. Sütçü and Ç. AytėkiN, “AN EXAMPLE OF PRAGMATIC ANALYSIS IN NATURAL LANGUAGE PROCESSING: SENTIMENTAL ANALYSIS OF MOVIE REVIEWS,” *Turk. ONLINE J. Des. ART Commun.*, vol. CTC, no. 2019, pp. 61–74, Dec. 2019, doi: 10.7456/ctc_2019_05.
- [49] A. Vaswani *et al.*, “Attention Is All You Need,” Aug. 01, 2023, *arXiv: arXiv:1706.03762*. Accessed: June 20, 2024. [Online]. Available: <http://arxiv.org/abs/1706.03762>
- [50] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, “Improving Language Understanding by Generative Pre-Training”.
- [51] P. J. Liu *et al.*, “Generating Wikipedia by Summarizing Long Sequences,” Jan. 30, 2018, *arXiv: arXiv:1801.10198*. doi: 10.48550/arXiv.1801.10198.
- [52] N. Kitaev and D. Klein, “Constituency Parsing with a Self-Attentive Encoder,” May 02, 2018, *arXiv: arXiv:1805.01052*. doi: 10.48550/arXiv.1805.01052.
- [53] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” May 24, 2019, *arXiv: arXiv:1810.04805*. doi: 10.48550/arXiv.1810.04805.
- [54] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, “Language Models are Unsupervised Multitask Learners”.
- [55] J. L. Ba, J. R. Kiros, and G. E. Hinton, “Layer Normalization,” July 21, 2016, *arXiv: arXiv:1607.06450*. doi: 10.48550/arXiv.1607.06450.
- [56] F. Hill, A. Bordes, S. Chopra, and J. Weston, “The Goldilocks Principle: Reading Children’s Books with Explicit Memory Representations,” Apr. 01, 2016, *arXiv: arXiv:1511.02301*. doi: 10.48550/arXiv.1511.02301.
- [57] D. Paperno *et al.*, “The LAMBADA dataset: Word prediction requiring a broad discourse context,” June 20, 2016, *arXiv: arXiv:1606.06031*. doi: 10.48550/arXiv.1606.06031.
- [58] H. Levesque, E. Davis, and L. Morgenstern, “The Winograd Schema Challenge,” 2012.
- [59] S. Reddy, D. Chen, and C. D. Manning, “CoQA: A Conversational Question

- Answering Challenge,” Mar. 29, 2019, *arXiv*: arXiv:1808.07042. doi: 10.48550/arXiv.1808.07042.
- [60] R. Nallapati, B. Zhou, C. N. dos santos, C. Gulcehre, and B. Xiang, “Abstractive Text Summarization Using Sequence-to-Sequence RNNs and Beyond,” Aug. 26, 2016, *arXiv*: arXiv:1602.06023. doi: 10.48550/arXiv.1602.06023.
- [61] M. E. Peters *et al.*, “Deep contextualized word representations,” Mar. 22, 2018, *arXiv*: arXiv:1802.05365. doi: 10.48550/arXiv.1802.05365.
- [62] Y. Jernite, S. R. Bowman, and D. Sontag, “Discourse-Based Objectives for Fast Unsupervised Sentence Representation Learning,” Apr. 23, 2017, *arXiv*: arXiv:1705.00557. doi: 10.48550/arXiv.1705.00557.
- [63] A. Wang, A. Singh, J. Michael, F. Hill, O. Levy, and S. R. Bowman, “GLUE: A Multi-Task Benchmark and Analysis Platform for Natural Language Understanding,” Feb. 22, 2019, *arXiv*: arXiv:1804.07461. doi: 10.48550/arXiv.1804.07461.
- [64] P. Rajpurkar, J. Zhang, K. Lopyrev, and P. Liang, “SQuAD: 100,000+ Questions for Machine Comprehension of Text,” Oct. 11, 2016, *arXiv*: arXiv:1606.05250. doi: 10.48550/arXiv.1606.05250.
- [65] R. Zellers, Y. Bisk, R. Schwartz, and Y. Choi, “SWAG: A Large-Scale Adversarial Dataset for Grounded Commonsense Inference,” Aug. 16, 2018, *arXiv*: arXiv:1808.05326. doi: 10.48550/arXiv.1808.05326.
- [66] P. Gage, “A new algorithm for data compression,” *C Users J*, vol. 12, no. 2, pp. 23–38, Feb. 1994.
- [67] I. Loshchilov and F. Hutter, “Decoupled Weight Decay Regularization,” Jan. 04, 2019, *arXiv*: arXiv:1711.05101. doi: 10.48550/arXiv.1711.05101.
- [68] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” Jan. 29, 2017, *arXiv*: arXiv:1412.6980. Accessed: Oct. 14, 2023. [Online]. Available: <http://arxiv.org/abs/1412.6980>
- [69] N. Mohseni, H. Nematzadeh, E. Akbarib, and H. Motameni, “Outlier Detection in Test Samples using Standard Deviation and Unsupervised Training Set Selection,” *Int. J. Eng.*, vol. 36, no. 1, pp. 119–129, 2023, doi: 10.5829/IJE.2023.36.01A.14.
- [70] J. Qi, J. Du, S. M. Siniscalchi, X. Ma, and C.-H. Lee, “On Mean Absolute Error for Deep Neural Network Based Vector-to-Vector Regression,” *IEEE Signal Process. Lett.*, vol. 27, pp. 1485–1489, 2020, doi: 10.1109/LSP.2020.3016837.
- [71] B. A. Kitchenham, L. M. Pickard, S. G. MacDonell, and M. J. Shepperd, “What accuracy statistics really measure,” *IEE Proc. - Softw.*, vol. 148, no. 3, p. 81, 2001, doi: 10.1049/ip-sen:20010506.
- [72] R. E. Walpole, R. H. Myers, J. DeShannon, and K. Ye, *Probability & statistics for engineers & scientists*, Ninth edition, Global edition. Boston Columbus Hoboken [und 22 weitere]: Pearson Education Limited, 2016.