

***Cross-Project Software Effort Estimation Pendekatan Transfer  
Learning Menggunakan Machine Learning dan Optuna  
Hyperparameter Optimization***



**TESIS**

Diajukan sebagai salah satu syarat untuk memperoleh gelar  
Magister Komputer (M.Kom)

**RIAN HIDAYAT**

**14230027**

**PROGRAM STUDI MAGISTER ILMU KOMPUTER (S2)**

**FAKULTAS TEKNOLOGI INFORMASI**

**UNIVERSITAS NUSA MANDIRI**

**JAKARTA**

**2025**

## SURAT PERNYATAAN ORISINALITAS DAN BEBAS PLAGIARISME

Yang bertanda tangan di bawah ini :

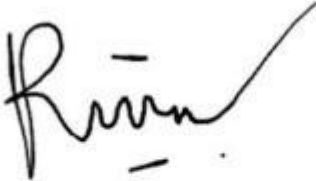
Nama : Rian Hidayat  
NIM : 14230027  
Program Studi : Ilmu Komputer  
Fakultas : Teknologi Informasi  
Jenjang : Strata Dua (S2)  
Peminatan : Software Engineering & Data Science

Dengan ini menyatakan bahwa tesis yang telah saya buat dengan judul: **“Cross-Project Software Effort Estimation Pendekatan Transfer Learning Menggunakan Machine Learning dan Optuna Hyperparameter Optimization”** adalah hasil karya sendiri, dan semua sumber baik yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar, serta belum pernah diterbitkan atau dipublikasikan dimanapun dan dalam bentuk apapun.

Demikianlah surat pernyataan ini saya buat dengan sebenar-benarnya. Apabila dikemudian hari ternyata saya memberikan keterangan palsu dan atau ada pihak lain yang mengklaim bahwa tesis yang telah saya buat adalah hasil karya milik seseorang atau badan tertentu, saya bersedia diproses baik secara pidana maupun perdata dan kelulusan saya dari Program Studi Ilmu Komputer (S2) Fakultas Teknologi Informasi Universitas Nusa Mandiri dicabut/dibatalkan.

Jakarta, 12 Februari 2026



  
Rian Hidayat

## SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan di bawah ini, saya :

Nama : Rian Hidayat  
NIM : 14230027  
Program Studi : Ilmu Komputer  
Fakultas : Teknologi Informasi  
Jenjang : Strata Dua (S2)  
Peminatan : Software Engineering & Data Science  
Jenis Karya : Tesis

Demi pengembangan ilmu pengetahuan, dengan ini menyetujui untuk memberikan izin kepada pihak Program Studi Ilmu Komputer (S2) Fakultas Teknologi Informasi Universitas Nusa Mandiri, **Hak Bebas Royalti Non-Eksklusif (Non-exclusive Royalti-Free Right)** atas karya ilmiah saya yang berjudul : **“Cross-Project Software Effort Estimation Pendekatan Transfer Learning Menggunakan Machine Learning dan Optuna Hyperparameter Optimization”** beserta perangkat yang diperlukan (apabila ada).

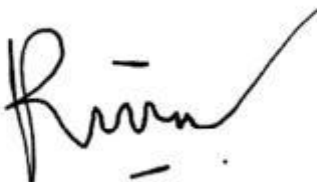
Dengan **Hak Bebas Royalti Non-Eksklusif** ini, pihak Universitas Nusa Mandiri berhak menyimpan, mengalih-media atau bentuk-kan, mengelolanya dalam pangkalan data (database), mendistribusikannya dan menampilkan atau mempublikasikannya di internet atau media lain untuk kepentingan akademis tanpa perlu meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta karya ilmiah tersebut.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Universitas Nusa Mandiri, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 12 Februari 2026



  
Rian Hidayat

## HALAMAN PERSETUJUAN DAN PENGESAHAN TESIS

Tesis ini diajukan oleh:

Nama : Rian Hidayat  
NIM : 14230027  
Program Studi : Ilmu Komputer  
Jenjang : Strata Dua (S2)  
Konsentrasi : *Software Engineering & Data Science*  
Judul Tesis : Cross-Project Software Effort Estimation Pendekatan Transfer Learning Menggunakan Machine Learning Dan Optuna Hyperparameter Optimization

Telah dipertahankan pada periode 2025-2 dihadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Magister Ilmu Komputer (M.Kom) pada Program Studi Ilmu Komputer (S2) Fakultas Teknologi Informasi Universitas Nusa Mandiri.

Jakarta, 19 Februari 2026

### PEMBIMBING TESIS

Pembimbing I : Prof. Dr. Eng.. Hilman F. Pardede,  
S.T.,M. Eng.



### DEWAN PENGUJI

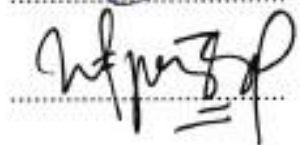
Penguji I : Prof. . Agus Subekti, M.T.



Penguji II : Dr. Kursehi Falgenti, M.Kom.



Penguji III /  
Pembimbing I : Prof. Dr. Eng.. Hilman F. Pardede,  
S.T.,M. Eng.



**LEMBAR BIMBINGAN TESIS****UNIVERSITAS NUSA MANDRI**

NIM : 14230027  
Nama Lengkap : Rian Hidayat  
Dosen Pembimbing : Prof. Dr. Eng. Hilman Ferdinandus Pardede, S.T., M.Eng.  
Judul Tesis : *Cross-Project Software Effort Estimation Pendekatan Transfer Learning Menggunakan Machine Learning dan Optuna Hyperparameter Optimization*

No	Tanggal Bimbingan	Pokok Bahasan	Paraf Dosen Pembimbing
1	16 September 2025	Pengajuan Judul dan Bab I	
2	02 Oktober 2025	Revisi Bab I dan Pengajuan Bab II serta Bab III	
3	03 Oktober 2025	Pemantapan Bab I, II dan III dalam slide Powerpoint	
4	07 Oktober 2025	Revisi Slide Present Seminar Proposal	
5	08 Oktober 2025	Revisi kedua Slide Present Seminar Proposal	
6	14 Oktober 2025	Laporan Hasil Eksperimen	
7	07 November 2025	Revisi Laporan Hasil Eksperimen	
8	03 Februari 2026	Acc keseluruhan Laporan Hasil Eksperimen dan Tulisan Bab IV dan V	

Catatan untuk dosen pembimbing  
Bimbingan Tesis

- \* Dimulai pada tanggal : 16 September 2025
- \* Diakhiri pada tanggal : 03 Februari 2026
- \* Jumlah pertemuan bimbingan : 8 (delapan) kali bimbingan

Disetujui oleh,  
Dosen Pembimbing

(Prof. Dr. Eng. Hilman Ferdinandus Pardede, S.T., M.Eng.)

## KATA PENGANTAR

Puji syukur alhamdulillah, penulis panjatkan kehadirat Allah, SWT, yang telah melimpahkan rahmat dan karunia-Nya, sehingga pada akhirnya penulis dapat menyelesaikan laporan tesis ini tepat pada waktunya. Dimana laporan tesis ini penulis sajikan dalam bentuk buku yang sederhana. Adapun judul tesis, yang penulis ambil sebagai berikut ***“Cross-Project Software Effort Estimation Pendekatan Transfer Learning Menggunakan Machine Learning dan Optuna Hyperparameter Optimization”***.

Tujuan penulisan laporan tesis ini dibuat sebagai salah satu syarat untuk Mendapatkan gelar Magister Komputer (M.Kom) pada Program Studi Ilmu Komputer (S2) Fakultas Teknologi Informasi Universitas Nusa Mandiri.

Penulis juga mencari dan menganalisa berbagai macam sumber referensi, baik dalam bentuk jurnal ilmiah, buku-buku literatur, internet, dll yang terkait dengan pembahasan pada laporan tesis ini.

Penulis menyadari bahwa tanpa bimbingan dan dukungan dari semua pihak dalam pembuatan laporan tesis ini, maka penulis tidak dapat menyelesaikan laporan tesis ini tepat pada waktunya. Untuk itu pada kesempatan ini penulis menyampaikan ucapan terimakasih kepada :

1. Rektor Universitas Nusa Mandiri
2. Wakil Rektor I Bidang Akademik Universitas Nusa Mandiri
3. Dekan Fakultas Teknologi Informasi Universitas Nusa Mandiri
4. Ketua Program Studi Ilmu Komputer (S2)
5. Bapak Prof. Dr. Eng. Hilman Ferdinandus Pardede, S.T., M.Eng. selaku dosen pembimbing yang telah menyediakan waktu, pikiran dan tenaga dalam membimbing penulis dalam menyelesaikan tesis ini
6. Orang tua tercinta yang telah memberikan dukungan material dan moral kepada penulis
7. Seluruh Dosen Program Studi Ilmu Komputer (S2) Fakultas Teknologi Informasi Universitas Nusa Mandiri yangtelah memberikan pelajaran yang berarti bagi penulis selama menempuh studi
8. Seluruh staff di lingkungan Universitas Nusa Mandiri yang telah melayani penulis dengan baik selama kuliah.

Serta semua pihak yang terlalu banyak untuk penulis sebutkan satu persatu sehingga terwujudnya penulisan laporan tesis ini. Penulis menyadari bahwa penulisan laporan tesis ini masih jauh sekali dari sempurna, untuk itu penulis mohon kritik dan saran yang bersifat membangun demi kesempurnaan penulisan karya ilmiah di masa yang akan datang.

Akhir kata semoga laporan tesis ini dapat bermanfaat bagi penulis khususnya dan bagi para pembaca yang berminat pada umumnya.



## ABSTRAK

Nama : Rian Hidayat  
NIM : 14230027  
Program Studi : Ilmu Komputer  
Fakultas : Teknologi Informasi  
Jenjang : Strata Dua (S2)  
Peminatan : Software Engineering & Data Science  
Judul Tesis : *Cross-Project Software Effort Estimation Pendekatan Transfer Learning Menggunakan Machine Learning dan Optuna Hyperparameter Optimization*

Perkiraan upaya pengembangan perangkat lunak yang akurat tetap menjadi tantangan kritis dalam rekayasa perangkat lunak, terutama dalam skenario lintas perusahaan atau lintas dataset di mana heterogenitas data dalam fitur, skala, dan distribusi menurunkan kinerja model prediksi tradisional. Studi ini menyelidiki penerapan kerangka kerja *Heterogeneous Transfer Learning* (HTL) berbasis *Deep Neural Network* (DNN) untuk meningkatkan akurasi perkiraan upaya di seluruh dataset heterogen.

Pendekatan yang diusulkan dimulai dengan identifikasi dan penyesuaian semantik fitur-fitur umum dari lima dataset yang tersedia secara publik: China, NASA93, Desharnais, Maxwell, dan SEERA. Pra-pemrosesan data meliputi transformasi Log1p untuk mengurangi kemiringan, pembersihan data, dan imputasi. Model DNN dasar dilatih pada dataset China dan dioptimalkan menggunakan Optuna untuk penyesuaian hiperparameter. Model yang dioptimalkan kemudian digunakan sebagai model sumber yang telah dilatih sebelumnya dalam eksperimen transfer learning, di mana berbagai strategi adaptasi dievaluasi dengan mengubah pembekuan dan pembukaan lapisan tersembunyi, mulai dari pembekuan penuh hingga penyesuaian penuh.

Hasil eksperimen menunjukkan bahwa efektivitas kerangka kerja HTL sangat bergantung pada kesesuaian antara strategi adaptasi dan karakteristik dataset target. Peningkatan kinerja yang signifikan diamati pada NASA93 ( $R^2$  meningkat dari 0,20 menjadi 0,56) dan SEERA ( $R^2$  mencapai 0,9173 pada *fine-tuning* penuh). Dataset Maxwell menunjukkan kompatibilitas tinggi, mencapai  $R^2$  0,85 bahkan dengan representasi yang dibekukan. Sebaliknya, Desharnais tetap menantang, dengan nilai  $R^2$  yang rendah secara konsisten (0,16), menunjukkan kesenjangan domain yang ekstrem. Temuan ini mengonfirmasi kelayakan transfer learning heterogen untuk estimasi upaya perangkat lunak lintas dataset dan menyoroti bahwa tidak ada strategi *fine-tuning* universal yang ada. Sebaliknya, kedalaman adaptasi harus disesuaikan dengan tingkat heterogenitas setiap *domain* target, memberikan wawasan praktis untuk mengembangkan model estimasi upaya yang lebih *robust* dan *generalisable*.

Kata kunci : *Software Effort Estimation; Heterogeneous Transfer Learning; Deep Neural Networks; Cross-Company Prediction; Domain Adaptation; Hyperparameter Optimization*

## ABSTRACT

Nama : Rian Hidayat  
NIM : 14230027  
Program Studi : Ilmu Komputer  
Fakultas : Teknologi Informasi  
Jenjang : Strata Dua (S2)  
Peminatan : Software Engineering & Data Science  
Judul Tesis : *Cross-Project Software Effort Estimation Pendekatan Transfer Learning Menggunakan Machine Learning dan Optuna Hyperparameter Optimization*

Accurate software development effort estimation remains a critical challenge in software engineering, particularly in cross-company or cross-dataset scenarios where data heterogeneity in features, scale, and distribution degrades the performance of traditional prediction models. This study investigates the application of a Heterogeneous Transfer Learning (HTL) framework based on Deep Neural Networks (DNN) to improve effort estimation accuracy across heterogeneous datasets.

The proposed approach begins with the identification and semantic alignment of common features from five publicly available datasets: China, NASA93, Desharnais, Maxwell, and SEERA. Data preprocessing includes Log1p transformation to reduce skewness, data cleaning, and imputation. A base DNN model is trained on the China dataset and optimized using Optuna for hyperparameter tuning. The optimized model is then employed as a pre-trained source model in transfer learning experiments, where multiple adaptation strategies are evaluated by varying the freezing and unfreezing of hidden layers, ranging from full freezing to full fine-tuning.

Experimental results demonstrate that the effectiveness of the HTL framework strongly depends on the compatibility between the adaptation strategy and the target dataset characteristics. Substantial performance improvements are observed on NASA93 ( $R^2$  increased from 0.20 to 0.56) and SEERA ( $R^2$  reaching 0.9173 under full fine-tuning). The Maxwell dataset exhibits high compatibility, achieving an  $R^2$  of 0.85 even with frozen representations. In contrast, Desharnais remains challenging, with persistently low  $R^2$  values (0.16), indicating an extreme domain gap.

These findings confirm the feasibility of heterogeneous transfer learning for cross-dataset software effort estimation and highlight that no universal fine-tuning strategy exists. Instead, the depth of adaptation must be tailored to the heterogeneity level of each target domain, providing practical insights for developing more robust and generalisable effort estimation models.

Kata kunci : Software Effort Estimation; Heterogeneous Transfer Learning; Deep Neural Networks; Cross-Company Prediction; Domain Adaptation; Hyperparameter Optimization

## DAFTAR REFERENSI

- [1] G. Silberschatz, A. Galvin, P. B., & Gagne, *Operating system concepts (10th ed.)*, 10th ed. Laurie Rosatone, 2018. [Online]. Available: <https://os.ecci.ucr.ac.cr/slides/Abraham-Silberschatz-Operating-System-Concepts-10th-2018.pdf>
- [2] A. Jadhav, S. K. Shandilya, I. Izonin, and R. Muzyka, “Multi-Step Dynamic Ensemble Selection to Estimate Software Effort,” *Appl. Artif. Intell.*, 2024, doi: 10.1080/08839514.2024.2351718.
- [3] K. Lavingia, V. Patel, and A. Lavingia, “Software Effort Estimation using Machine Learning Algorithms,” *Scalable Comput. Pract. Exp.*, vol. 25, no. 2, pp. 1276–1285, 2024, doi: 10.12694/scpe.v25i2.2213.
- [4] T. M. K. Kumar and M. A. Jayaram, “Software Effort Estimation Using Hard Limiting Techniques with Special Reference to Small Size Technical & Analytical Projects,” in *International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT)*, Mandya, India: IEEE Fourth International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT), 2022, pp. 1–8. doi: 10.1109/ICERECT56837.2022.10060696.
- [5] D. M. Singh and D. R. Tripathi, “Traditional and recent techniques of effort estimation in software engineering,” *BSSS J. Comput.*, 2023, doi: 10.51767/jc1403.
- [6] farah alhamdany and L. Ibrahim, “Software Development Effort Estimation Techniques: A Survey,” *مجلة التربية والعلم*, vol. 31, no. 1, pp. 80–92, 2022, doi: 10.33899/edusj.2022.132274.1201.
- [7] M. Jorgensen and M. Shepperd, “A Systematic Review of Software Development Cost Estimation Studies,” *IEEE Trans. Softw. Eng.*, vol. 33, no. 1, pp. 33–53, 2007, doi: 10.1109/TSE.2007.256943.
- [8] J. Wen, S. Li, Z. Lin, Y. Hu, and C. Huang, “Systematic literature review of machine learning based software development effort estimation models,” *Inf. Softw. Technol.*, vol. 54, no. 1, pp. 41–59, Jan. 2012, doi: 10.1016/j.infsof.2011.09.002.
- [9] S. W. I. Kuan, “Factors on Software Effort Estimation,” *Int. J. Softw. Eng.*

- Appl.*, vol. 8, no. 1, pp. 23–32, 2017, doi: 10.5121/IJSEA.2017.8103.
- [10] M. Padmaja and D. Haritha, “Software Effort Estimation Using Grey Relational Analysis,” *Int. J. Inf. Technol. Comput. Sci.*, vol. 9, no. 5, pp. 52–60, 2017, doi: 10.5815/IJITCS.2017.05.07.
- [11] E. Kocaguneli, T. Menzies, and E. Mendes, “Transfer learning in effort estimation,” *Empir. Softw. Eng.*, vol. 20, no. 3, pp. 813–843, 2015, doi: 10.1007/S10664-014-9300-5.
- [12] S. Kumari, R. . C, S. K, K. Selvi, T. A. Mohanaprakash, and C. Tamilselvi, “Optuna-Optimized Machine Learning Technique for Accurate Diabetes Prediction and Classification,” in *2024 4th International Conference on Sustainable Expert Systems (ICSES)*, IEEE, Oct. 2024, pp. 1478–1485. doi: 10.1109/ICSES63445.2024.10763036.
- [13] B. W. Boehm, “Software Engineering Economics,” *IEEE Trans. Softw. Eng.*, vol. SE-10, no. 1, pp. 4–21, Jan. 1984, doi: 10.1109/TSE.1984.5010193.
- [14] B. Kitchenham and E. Mendes, “Why comparative effort prediction studies may be invalid,” in *Proceedings of the 5th International Conference on Predictor Models in Software Engineering*, New York, NY, USA: ACM, May 2009, pp. 1–5. doi: 10.1145/1540438.1540444.
- [15] L. L. Minku and S. Hou, “Clustering Dycom,” in *Proceedings of the 13th International Conference on Predictive Models and Data Analytics in Software Engineering*, New York, NY, USA: ACM, Nov. 2017, pp. 12–21. doi: 10.1145/3127005.3127007.
- [16] F. Ferrucci, E. Mendes, and F. Sarro, “Web effort estimation,” in *Proceedings of the 8th International Conference on Predictive Models in Software Engineering*, New York, NY, USA: ACM, Sep. 2012, pp. 29–38. doi: 10.1145/2365324.2365330.
- [17] N. Papernot, M. Abadi, Ú. Erlingsson, I. Goodfellow, and K. Talwar, “Semi-supervised knowledge transfer for deep learning from private training data,” in *International Conference on Learning Representations*, 2018. [Online]. Available: <https://openreview.net/pdf?id=HkY2BZcee>
- [18] C. M. . Bishop, *Pattern recognition and machine learning*. Springer, 2016.
- [19] S. J. Pan and Q. Yang, “A Survey on Transfer Learning,” *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010, doi:

- 10.1109/TKDE.2009.191.
- [20] M. C. Weiss *et al.*, “The physiology and habitat of the last universal common ancestor,” *Nat. Microbiol.*, vol. 1, no. 9, p. 16116, Jul. 2016, doi: 10.1038/nmicrobiol.2016.116.
- [21] S. Tariq, M. Usman, R. Wong, Y. Zhuang, and S. Fong, “On Learning Software Effort Estimation,” in *2015 3rd International Symposium on Computational and Business Intelligence (ISCBI)*, IEEE, Dec. 2015, pp. 79–84. doi: 10.1109/ISCBI.2015.21.
- [22] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, “How transferable are features in deep neural networks?,” Nov. 2014.
- [23] C. Zhuang *et al.*, “Unsupervised neural network models of the ventral visual stream,” *Proc. Natl. Acad. Sci.*, vol. 118, no. 3, Jan. 2021, doi: 10.1073/pnas.2014196118.
- [24] M. Abadi *et al.*, “Deep learning with differential privacy,” *Commun. ACM*, vol. 65, no. 10, pp. 61–71, 2022, doi: 10.1145/3549522.
- [25] J. Chung and J. Teo, “Single classifier vs. ensemble machine learning approaches for mental health prediction,” *Brain Informatics*, vol. 10, no. 1, pp. 1–10, 2023, doi: 10.1186/s40708-022-00180-6.
- [26] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, “A Survey on Deep Transfer Learning,” Aug. 2018.
- [27] A. S. Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, “CNN Features off-the-shelf: an Astounding Baseline for Recognition,” May 2014.
- [28] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, “Optuna: A Next-generation Hyperparameter Optimization Framework,” *arXiv: Learning*. 2019.
- [29] Optuna developers, “Optuna documentation.” Accessed: Nov. 29, 2023. [Online]. Available: <https://optuna.readthedocs.io/>
- [30] B. Kiran Kumar, M. Rekha Sundari, and A. Surekha, “A Deep Learning-Based Technique for Evaluation of Estimation in Software Development,” 2024, pp. 291–306. doi: 10.1007/978-981-97-6995-7\_22.
- [31] J. W. Osborne, “Improving your data transformations: Applying the Box-Cox transformation,” *Pract. Assessment, Res. Eval.*, vol. 15, no. 12, 2010.
- [32] A. Dorado *et al.*, “A Process Mining-Based System For The Analysis and Prediction of Software Development Workflows,” Oct. 2025.

- [33] C. Hu *et al.*, “Scalable imaging-free spatial genomics through computational reconstruction.” Aug. 07, 2024. doi: 10.1101/2024.08.05.606465.
- [34] & J. K. Kuhn M., “Feature Engineering and Selection: A Practical Approach for Predictive Models.,” *Chapman Hall*, 2020.
- [35] P. A. Whigham, C. A. Owen, and S. G. MacDonell, “A Baseline Model for Software Effort Estimation,” *arXiv Softw. Eng.*, 2021, doi: 10.1145/2738037.
- [36] Y. Assefa, F. Berhanu, A. Tilahun, and E. Alemneh, “Software Effort Estimation using Machine learning Algorithm,” in *2022 International Conference on Information and Communication Technology for Development for Africa (ICT4DA)*, IEEE, Nov. 2022, pp. 163–168. doi: 10.1109/ICT4DA56482.2022.9971209.
- [37] E. I. Mustafa and R. Osman, “SEERA: a software cost estimation dataset for constrained environments,” in *Proceedings of the 16th ACM International Conference on Predictive Models and Data Analytics in Software Engineering*, New York, NY, USA: ACM, Nov. 2020, pp. 61–70. doi: 10.1145/3416508.3417119.
- [38] Adnan Purwanto, Lindung Parningotan Manik, “Software Effort Estimation Using Logarithmic Fuzzy Preference Programming and Least Squares Support Vector Machines.pdf,” *Sci. J. Informatics*, vol. 10, no. 1, pp. 1–12, 2023, [Online]. Available: <https://pdfs.semanticscholar.org/db16/8396ba9c085705ea2df94f6d08e64435eb91.pdf>
- [39] L. P. Manik, “Features Engineering with Generating Genetic Algorithm for Software Effort Estimation,” Jan. 2022, doi: 10.24507/icicelb.13.01.1.
- [40] H. R. Moechtar, S. Sulistyono, and R. Ferdiana, “A Review on Improving the Accuracy of Effort Estimation in Software Development with Agile Method,” in *2024 11th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, 2024, pp. 31–36. doi: 10.1109/ICITACEE62763.2024.10761956.
- [41] L. Cao, “Estimating Efforts for Various Activities in Agile Software Development: An Empirical Study,” *IEEE Access*, vol. 10, pp. 83311–83321, 2022, doi: 10.1109/ACCESS.2022.3196923.
- [42] S. Wakurdekar, S. Vanjale, P. Paygude, M. D. Gayakwad, A. Kadam, and R.

- Joshi, "Novel Approach to Design a Model for Software Effort Estimation Using Linear Regression," *J. Electr. Syst.*, 2024, doi: 10.52783/jes.1997.
- [43] M. Awais and A. A. Malik, "A Comparative Analysis of Popular Software Effort Estimation Techniques," in *2023 International Conference on IT and Industrial Technologies (ICIT)*, IEEE, Oct. 2023, pp. 1–5. doi: 10.1109/ICIT59216.2023.10335868.
- [44] Z. P. A. P. S. HUYNH THAI HOC, RADEK SILHAVY, "Comparing Stacking Ensemble and Deep Learning for Software Project Effort Estimation," *IEEE Access*, vol. 11, pp. 60590–60604, 2023, doi: 10.1109/access.2023.3286372.
- [45] S. Amasaki, "Augmenting Window Contents with Transfer Learning for Effort Estimation.," in *CEUR Workshop Proceedings*, 2020, pp. 4–12.
- [46] B. Setiawan and A. Subekti, "Multi View Natural Network for Cross- Project Software Defect Prediction," no. April, pp. 41–53, 2025, doi: 10.52985/insyst.v7i1.436.
- [47] A. S. Kirso, "ENSEMBLE AND VOTING APPROACHES FOR DEFECT PREDICTION Universitas Nusa Mandiri Cross project defect prediction ( CPDP ) has emerged as a critical area within software engineering , offering a promising approach for enhancing software quality and maintainab," *MUST J. Math. Educ. Sci. Technol.*, vol. 10, no. 1, pp. 58–67, 2025.
- [48] S. S. Ali, J. Ren, and J. Wu, "Framework to improve software effort estimation accuracy using novel ensemble rule," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 36, no. 9, p. 102189, Nov. 2024, doi: 10.1016/j.jksuci.2024.102189.
- [49] K. Korenaga, A. Monden, and Z. Yucel, "Data Smoothing for Software Effort Estimation," in *2019 20th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*, IEEE, Jul. 2019, pp. 501–506. doi: 10.1109/SNPD.2019.8935841.
- [50] S. S. Gautam and V. Singh, "Adaptive Discretization Using Golden Section to Aid Outlier Detection for Software Development Effort Estimation," *IEEE Access*, vol. 10, pp. 90369–90387, 2022, doi: 10.1109/ACCESS.2022.3200149.

- [51] I. Myrtveit, E. Stensrud, and U. H. Olsson, “Analyzing data sets with missing data: an empirical evaluation of imputation methods and likelihood-based methods,” *IEEE Trans. Softw. Eng.*, vol. 27, no. 11, pp. 999–1013, 2001, doi: 10.1109/32.965340.
- [52] B. Twala and M. Cartwright, “Ensemble missing data techniques for software effort prediction,” *Intell. Data Anal.*, vol. 14, no. 3, pp. 299–331, May 2010, doi: 10.3233/IDA-2010-0423.



## DAFTAR RIWAYAT HIDUP

### A. Biodata Mahasiswa

NIM : 14230027  
Nama Lengkap : Rian Hidayat  
Tempat dan Tanggal Lahir : Bogor, 26 Mei 1993  
Jenis Kelamin : Laki-laki  
Alamat : Jl. H. Nurdin No. 79 RT/RW 002/001 Kel.  
Cipayung Kec. Cipayung Jaya Kota Depok

### B. Riwayat Pendidikan


1. MI Muhammadiyah 2 Depok, Lulus Tahun 2005
2. SMP Muhammadiyah 3 Depok, Lulus Tahun 2008
3. SMK Almuhtadin Depok, Lulus Tahun 2011
4. S1 Teknik Informatika Universitas MH Thamrin, Lulus Tahun 2020

### C. Riwayat Pengalaman Pekerjaan

1. IT Support STIKes Indonesia Maju Tahun 2012 - 2018
2. IT Support Politeknik Karya Husada Tahun 2019 - Sekarang



Jakarta, 12 Februari 2026

  
Rian Hidayat

UNIVERSITAS  
NUSA MANDIRI

## DAFTAR LAMPIRAN

### Lampiran 1 Deskripsi Feature Dataset China

Features	Deskripsi
AFP	Adjusted Function Points, yaitu ukuran perangkat lunak yang telah disesuaikan dengan faktor kompleksitas sistem menggunakan <i>Value Adjustment Factor</i> .
Input	Jumlah <i>External Input (EI)</i> dalam Function Point, yaitu proses yang menerima data dari pengguna atau sistem eksternal ke dalam sistem.
Output	Jumlah <i>External Output (EO)</i> dalam Function Point, yaitu proses yang menghasilkan keluaran berupa laporan atau informasi dari sistem.
Enquiry	Jumlah <i>External Inquiry (EQ)</i> , yaitu permintaan informasi yang hanya mengambil data tanpa pemrosesan kompleks.
File	Jumlah <i>Internal Logical Files (ILF)</i> , yaitu struktur data atau tabel yang dikelola oleh sistem.
Interface	Jumlah <i>External Interface Files (EIF)</i> yang digunakan oleh sistem tetapi dikelola oleh sistem lain.
Added	Jumlah fungsi baru yang ditambahkan selama proses pengembangan perangkat lunak.
Changed	Jumlah fungsi yang mengalami perubahan atau modifikasi selama pengembangan.
Deleted	Jumlah fungsi yang dihapus selama proses pengembangan perangkat lunak.
PDR_AFP	Productivity Delivery Rate berdasarkan Adjusted Function Points, yang menunjukkan tingkat produktivitas tim pengembang dalam menghasilkan perangkat lunak per unit effort.
PDR_UFP	Productivity Delivery Rate berdasarkan Unadjusted Function Points (UFP), yaitu ukuran produktivitas sebelum dilakukan penyesuaian kompleksitas.
NPDR_AFP	Normalized Productivity Delivery Rate berbasis AFP, yaitu nilai produktivitas yang telah dinormalisasi untuk mengurangi variasi antar proyek.
NPDU_UFP	Normalized Productivity Delivery Rate berbasis UFP yang telah dinormalisasi untuk meningkatkan konsistensi perbandingan antar proyek.
Resource	Menunjukkan jenis atau kategori tim pengembang yang digunakan dalam proyek perangkat lunak.
Dev.Type	Menunjukkan tipe pengembangan perangkat lunak, misalnya pengembangan baru, pemeliharaan, atau peningkatan sistem.

Duration	Total waktu yang dibutuhkan untuk menyelesaikan proyek perangkat lunak dari awal hingga selesai.
N_effort	Effort yang telah dinormalisasi untuk mengurangi variasi data antar proyek.
Effort	Total usaha atau tenaga kerja yang digunakan untuk menyelesaikan proyek perangkat lunak, biasanya dalam satuan person-month atau person-hour.



Lampiran 2 Deskripsi Feature Dataset Nasa93

Features	Deskripsi
Rely	Tingkat kebutuhan reliabilitas perangkat lunak. Semakin tinggi nilainya, semakin tinggi tuntutan sistem untuk beroperasi secara stabil tanpa kegagalan.
Data	Ukuran basis data yang digunakan oleh aplikasi. Menggambarkan volume data yang harus dikelola oleh sistem.
Cplx	Tingkat kompleksitas produk perangkat lunak yang dikembangkan, termasuk kompleksitas algoritma, struktur program, dan logika sistem.
Time	Batasan performa run-time yang diperlukan oleh sistem, misalnya kebutuhan respon cepat atau performa tinggi saat eksekusi.
Stor	Batasan penggunaan memori pada sistem selama eksekusi program.
Virt	Tingkat volatilitas lingkungan mesin virtual atau sistem operasi tempat perangkat lunak dijalankan.
Turn	Waktu yang dibutuhkan untuk melakukan proses komputasi atau <i>turnaround time</i> , yaitu waktu respon sistem terhadap permintaan pengguna.
Acap	Kemampuan analis sistem dalam memahami kebutuhan pengguna dan merancang sistem.
Aexp	Pengalaman analis dalam domain aplikasi yang sedang dikembangkan.
Pcap	Kemampuan programmer dalam mengimplementasikan perangkat lunak secara efektif.
Vexp	Pengalaman tim pengembang dalam menggunakan lingkungan mesin virtual atau platform sistem yang digunakan.
Lexp	Pengalaman tim pengembang dalam menggunakan bahasa pemrograman yang digunakan dalam proyek.
Modp	Tingkat penerapan metode rekayasa perangkat lunak dalam proses pengembangan, seperti penggunaan metodologi pengembangan yang terstruktur.
Tool	Tingkat penggunaan alat bantu perangkat lunak seperti IDE, CASE tools, atau alat otomatisasi dalam pengembangan.
Sced	Batasan jadwal pengembangan proyek yang harus dipenuhi oleh tim pengembang.
equivphyskloc	Ukuran perangkat lunak dalam satuan ribuan baris kode ( <i>thousand lines of code / KLOC</i> ). Digunakan sebagai ukuran ukuran sistem.
Effort	Total usaha yang digunakan untuk menyelesaikan proyek perangkat lunak, biasanya diukur dalam satuan person-month. Ini merupakan variabel target dalam model estimasi effort.

Lampiran 3 Deskripsi Feature Dataset Desharnais

Features	Deskripsi
Project	Nomor atau identitas unik proyek perangkat lunak dalam dataset. Digunakan hanya sebagai identifier proyek.
TeamExp	Pengalaman rata-rata tim pengembang dalam satuan tahun. Nilai ini menunjukkan tingkat pengalaman tim dalam pengembangan perangkat lunak.
ManagerExp	Pengalaman manajer proyek dalam satuan tahun dalam mengelola proyek perangkat lunak.
YearEnd	Tahun penyelesaian proyek perangkat lunak. Variabel ini menggambarkan kapan proyek tersebut selesai dikembangkan.
Length	Durasi atau lama waktu pengerjaan proyek perangkat lunak dari awal hingga selesai.
Effort	Total usaha yang dikeluarkan untuk menyelesaikan proyek perangkat lunak, biasanya diukur dalam person-hour atau person-month. Variabel ini merupakan target variable dalam penelitian software effort estimation.
Transaction	Jumlah transaksi atau proses bisnis yang ditangani oleh sistem perangkat lunak.
Entities	Jumlah entitas data yang terdapat dalam sistem, biasanya merepresentasikan jumlah tabel atau objek data dalam database.
PointsNonAdjust	<i>Unadjusted Function Points (UFP)</i> , yaitu ukuran perangkat lunak berdasarkan jumlah fungsi tanpa mempertimbangkan faktor kompleksitas sistem.
Adjustment	Faktor penyesuaian yang digunakan untuk menyesuaikan ukuran Function Point berdasarkan kompleksitas sistem dan karakteristik teknis lainnya.
PointsAdjust	<i>Adjusted Function Points (AFP)</i> , yaitu ukuran perangkat lunak yang telah disesuaikan menggunakan faktor penyesuaian dari UFP.
Language	Bahasa pemrograman yang digunakan dalam pengembangan perangkat lunak, misalnya COBOL, C, atau bahasa lainnya.

Lampiran 4 Deskripsi Feature Dataset Maxwell

Features	Deskripsi
Year	Tahun dimulainya proyek pengembangan perangkat lunak.
Har	Platform perangkat keras ( <i>hardware platform</i> ) yang digunakan untuk mengembangkan atau menjalankan aplikasi.
App	Jenis atau tipe aplikasi yang sedang dikembangkan dalam proyek tersebut.
Db	Sistem manajemen basis data ( <i>Database Management System / DBMS</i> ) yang digunakan dalam proyek.
Ifc	Teknologi antarmuka pengguna ( <i>user interface technology</i> ) yang digunakan dalam aplikasi.
Source	Sistem manajemen kode sumber ( <i>source code management system</i> ) yang digunakan selama pengembangan proyek.
Telonus	Menunjukkan apakah proyek menggunakan IBM Telon, yaitu alat pengembangan legacy untuk sistem mainframe.
Nlan	Jumlah bahasa pemrograman yang digunakan dalam proyek perangkat lunak.
T01	Tingkat partisipasi pelanggan dalam proses pengembangan perangkat lunak.
T02	Kecukupan atau kesiapan lingkungan pengembangan perangkat lunak.
T03	Ketersediaan staf atau sumber daya manusia dalam proyek.
T04	Tingkat penggunaan standar pengembangan perangkat lunak dalam proyek.
T05	Tingkat penggunaan metode atau metodologi rekayasa perangkat lunak.
T06	Tingkat penggunaan alat bantu perangkat lunak ( <i>software tools</i> ) dalam pengembangan.
T07	Kompleksitas logika perangkat lunak yang dikembangkan.
T08	Tingkat perubahan atau volatilitas kebutuhan ( <i>requirements volatility</i> ) selama pengembangan.
T09	Tingkat kebutuhan kualitas perangkat lunak yang harus dipenuhi.

T10	Kebutuhan efisiensi sistem, seperti performa dan penggunaan sumber daya.
T11	Kebutuhan instalasi perangkat lunak pada lingkungan pengguna.
T12	Kemampuan staf dalam melakukan analisis sistem dan kebutuhan perangkat lunak.
T13	Pengetahuan staf terhadap domain aplikasi yang dikembangkan.
T14	Kemampuan staf dalam menggunakan alat bantu pengembangan perangkat lunak.
T15	Kemampuan kerja sama tim ( <i>teamwork skills</i> ) dari staf pengembang.
Duration	Durasi proyek dalam satuan bulan dari awal hingga selesai.
Size	Ukuran perangkat lunak yang dikembangkan dalam satuan Lines of Code (LOC).
Time	Total waktu yang dihabiskan untuk proyek dalam satuan person-month.
Effort	Total usaha atau tenaga kerja yang digunakan untuk menyelesaikan proyek dalam person-month. Variabel ini biasanya digunakan sebagai target variable dalam model estimasi effort.

UNIVERSITAS  
NUSA MANDIRI

Lampiran 5 Deskripsi Feature Dataset SEERA

Kategori	Features	Deskripsi
General Information	Project ID	Identitas unik proyek perangkat lunak
	Year of project	Tahun proyek dilakukan
	Organization ID	Identitas organisasi pengembang
	Organization type	Jenis organisasi pengembang
	Role in organization	Peran organisasi dalam proyek
	Size of organization	Ukuran organisasi pengembang
	Size of IT department	Ukuran departemen IT dalam organisasi
	Customer organization type	Jenis organisasi pelanggan
	Estimated duration	Perkiraan durasi proyek
	Actual duration	Durasi aktual proyek
	% project gain (loss)	Persentase keuntungan atau kerugian proyek
	Development type	Jenis pengembangan perangkat lunak
	Application Domain	Domain aplikasi yang dikembangkan
Contract	Contract date	Tanggal kontrak proyek dimulai
	Contract software delivery date	Tanggal target penyerahan perangkat lunak
	Actual software development start date	Tanggal aktual dimulainya pengembangan
Effort	Estimated Effort	Estimasi usaha pengembangan
	Actual Effort	Usaha aktual yang dikeluarkan
Cost	Contract price	Harga kontrak proyek
	Actual incurred costs	Biaya aktual yang dikeluarkan
Size	Object Points	Ukuran perangkat lunak menggunakan metode Object Points
	Other sizing method	Metode pengukuran ukuran perangkat lunak lainnya
	Estimated Size	Estimasi ukuran perangkat lunak
	Number of screens	Jumlah layar dalam aplikasi
	Number of reports	Jumlah laporan yang dihasilkan sistem
Environment	Contract maturity	Tingkat kematangan kontrak proyek
	Government policy impact	Pengaruh kebijakan pemerintah terhadap proyek

	Economic instability impact	Dampak ketidakstabilan ekonomi
	Organization management structure clarity	Kejelasan struktur manajemen organisasi
	Developer hiring policy	Kebijakan perekrutan developer
	Developer incentives policy	Kebijakan insentif untuk developer
	Developer training	Program pelatihan developer
	Development team management	Manajemen tim pengembangan
Users	Top management support	Dukungan manajemen tingkat atas
	Top management opinion of previous system	Pendapat manajemen tentang sistem sebelumnya
	Clarity of manual system	Kejelasan sistem manual sebelumnya
	User resistance	Tingkat resistensi pengguna terhadap sistem
	User computer experience	Pengalaman pengguna menggunakan komputer
	Users stability	Stabilitas pengguna sistem
	Requirements stability	Stabilitas kebutuhan sistem
	Requirements flexibility	Fleksibilitas perubahan kebutuhan
Developers	Project manager experience	Pengalaman manajer proyek
	Consultant availability	Ketersediaan konsultan
	DBMS expert availability	Ketersediaan ahli database
	Precedentedness	Tingkat kesamaan proyek dengan proyek sebelumnya
	Software tool experience	Pengalaman menggunakan tools pengembangan
	Programmers experience in language	Pengalaman programmer dalam bahasa pemrograman
	Programmers capability	Kemampuan programmer
	Analysts capability	Kemampuan analis sistem
	Team selection	Proses pemilihan anggota tim
	Team size	Jumlah anggota tim
	Dedicated team members	Anggota tim yang fokus pada proyek
	Daily working hours	Jam kerja harian tim
	Team contracts	Kontrak kerja tim
	Team continuity	Stabilitas anggota tim
	Team cohesion	Kekompakan tim

	Income satisfaction	Kepuasan anggota tim terhadap penghasilan
Product	Requirement accuracy level	Tingkat akurasi kebutuhan sistem
	Technical documentation	Ketersediaan dokumentasi teknis
	Comments within the code	Dokumentasi komentar dalam kode
	User manual	Dokumentasi panduan pengguna
	Required reusability	Kebutuhan penggunaan kembali komponen
	Performance requirements	Kebutuhan performa sistem
	Product complexity	Kompleksitas produk perangkat lunak
	Security requirements	Kebutuhan keamanan sistem
	Reliability requirements	Kebutuhan reliabilitas perangkat lunak
	Specified hardware	Spesifikasi perangkat keras yang dibutuhkan
Project	Schedule quality	Kualitas perencanaan jadwal proyek
	Development environment adequacy	Kecukupan lingkungan pengembangan
	Tool availability	Ketersediaan alat bantu pengembangan
	Methodology	Metodologi pengembangan yang digunakan
	Multiple programming languages	Penggunaan lebih dari satu bahasa pemrograman
	Programming language used	Bahasa pemrograman utama yang digunakan
	DBMS used	Sistem manajemen database yang digunakan
	Technical stability	Stabilitas teknis sistem
	Open source software	Penggunaan perangkat lunak open source
	Level of outsourcing	Tingkat penggunaan outsourcing
	Outsourcing impact	Dampak outsourcing terhadap proyek
	Degree of software reuse	Tingkat penggunaan kembali komponen perangkat lunak
	Degree of risk management	Tingkat manajemen risiko proyek
	Use of standards	Penggunaan standar pengembangan

	Degree of standards usage	Tingkat penerapan standar
	Process reengineering	Penerapan rekayasa ulang proses bisnis



## Lampiran 6 Bukti Hasil Pengecekan Plagiarisme Laporan Tesis

### DRAF TESIS

#### ORIGINALITY REPORT

<b>11%</b>	<b>10%</b>	<b>4%</b>	<b>9%</b>
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

#### PRIMARY SOURCES

<b>1</b>	<b>Submitted to Universitas Raharja</b> Student Paper	<b>5%</b>
<b>2</b>	<b>repository.nusamandiri.ac.id</b> Internet Source	<b>3%</b>
<b>3</b>	<b>repository.bsi.ac.id</b> Internet Source	<b>&lt;1%</b>
<b>4</b>	<b>elibrary.nusamandiri.ac.id</b> Internet Source	<b>&lt;1%</b>
<b>5</b>	<b>Submitted to Sekolah Teknik Elektro &amp; Informatika</b> Student Paper	<b>&lt;1%</b>
<b>6</b>	<b>journal.esrgroups.org</b> Internet Source	<b>&lt;1%</b>
<b>7</b>	<b>Haryadi Tri Nugroho, Syarif Hidayat. "Implementasi Intelejen Bisnis dengan Visualisasi Data Gaji dan Algoritma Linear Regresion", Jurnal Indonesia : Manajemen Informatika dan Komunikasi, 2024</b> Publication	<b>&lt;1%</b>
<b>8</b>	<b>doaj.org</b> Internet Source	<b>&lt;1%</b>
<b>9</b>	<b>jurnal.polibatam.ac.id</b> Internet Source	<b>&lt;1%</b>

## DAFTAR PUBLIKASI YANG TELAH DIHASILKAN

1. Hidayat, R., & Gata, W. (2025). Pemanfaatan IndoBERT Untuk Analisis Sentimen Ulasan Aplikasi Depok Single Window. *INFORMATION SYSTEM FOR EDUCATORS AND PROFESSIONALS : Journal Of Information System*, 10(1), 13-24. doi:[10.51211/isbi.v10i1.3334](https://doi.org/10.51211/isbi.v10i1.3334), url: <https://ejournal-binainsani.ac.id/index.php/ISBI/article/view/3334>
2. Hidayat, R., & Subekti, A. (2025). Evaluasi Kinerja Model Machine Learning dalam Cross-Project Defect Prediction Menggunakan Library PyCaret. *Techné : Jurnal Ilmiah Elektroteknika*, 24(2), 163–176. <https://doi.org/10.31358/techne.v24i2.603>, url: <https://ojs.jurnaltechne.org/index.php/techne/article/view/603>



UNIVERSITAS  
NUSA MANDIRI