

Stay or Leave? Predicting Employee Retention With Hybrid Deep Learning Models

Yessica Fara Desvia^{1*}, Wafiqah Yasmin Azhar², Supriyadi³, J. Rolles Herwin Sihombing⁴, Nindy Faoziyah⁵

^{1,2,3,4,5}Informatika, Universitas Horizon Indonesia, Jl. Pangkal Perjuangan By Pass No.KM.1, Tanjungpura, Kec. Karawang Bar., Karawang, Jawa Barat
E-mail: yessica.desvia.krw@horizon.ac.id

* Corresponding Author

 <https://doi.org/10.31004/jerkin.v4i4.6345>

ARTICLE INFO

Article history

Received: 15 Mar 2026

Revised: 23 April 2026

Accepted: 11 May 2026

Kata Kunci:

Retensi Karyawan
(Employee Retention),
Hybrid Deep Learning,
XGBoost dan Deep
Neural Network (DNN),
Prediksi Turnover
Karyawan, SHAP
Explainability

Keywords:

Employee Retention,
Hybrid Deep Learning,
XGBoost and Deep
Neural Network (DNN),
Employee Turnover
Prediction, SHAP
Explainability



ABSTRACT

Retensi karyawan menjadi tantangan utama organisasi di era digital karena tingginya turnover berdampak pada produktivitas, biaya operasional, dan performa organisasi. Penelitian ini mengusulkan model Hybrid Deep Learning berbasis XGBoost dan Deep Neural Network (DNN) untuk memprediksi retensi karyawan menggunakan dataset HR_comma_sep. Pendekatan ini menggabungkan machine learning berbasis tree dan deep learning untuk menangkap hubungan nonlinear serta pola keputusan kompleks. Preprocessing data dilakukan melalui feature scaling dan categorical encoding sebelum pelatihan model. Arsitektur hybrid dibangun dengan mengintegrasikan output probabilitas XGBoost dan DNN pada lapisan meta-classification. Evaluasi menggunakan Accuracy, Precision, Recall, F1-Score, dan AUC-ROC menunjukkan bahwa model hybrid memiliki performa prediksi dan generalisasi lebih baik dibandingkan metode konvensional. SHAP Explainability digunakan untuk mengidentifikasi faktor utama yang memengaruhi turnover, yaitu kepuasan kerja, rata-rata jam kerja bulanan, dan lama bekerja. Model ini dapat membantu organisasi menyusun strategi manajemen SDM secara proaktif.

Employee retention is a major challenge for organizations in the digital era because high turnover impacts productivity, operational costs, and organizational performance. This study proposes a Hybrid Deep Learning model based on XGBoost and Deep Neural Network (DNN) to predict employee retention using the HR_comma_sep dataset. This approach combines tree-based machine learning and deep learning to capture nonlinear relationships and complex decision patterns. Data preprocessing is performed through feature scaling and categorical encoding before model training. The hybrid architecture is built by integrating the probability outputs of XGBoost and DNN in the meta-classification layer. Evaluation using Accuracy, Precision, Recall, F1-Score, and AUC-ROC shows that the hybrid model has better prediction and generalization performance than conventional methods. SHAP Explainability is used to identify the main factors influencing turnover, namely job satisfaction, average monthly working hours, and length of service. This model can help organizations develop proactive HR management strategies.



This is an open access article under the CC-BY-SA license.

How to Cite: Yessica Fara Desvia, et al (2026). Stay or Leave? Predicting Employee Retention With Hybrid Deep Learning Models, 4(4) 25288-25295. <https://doi.org/10.31004/jerkin.v4i4.6345>

PENDAHULUAN

Retensi karyawan merupakan salah satu aspek penting dalam pengelolaan sumber daya manusia karena berkaitan langsung dengan stabilitas organisasi, produktivitas, dan efisiensi operasional. Tingginya tingkat turnover karyawan dapat menyebabkan peningkatan biaya rekrutmen, pelatihan, dan penurunan kualitas kinerja perusahaan (Ahmed & Yang, 2021; Hassan et al., 2021). Oleh karena itu,

organisasi memerlukan sistem prediksi yang mampu mengidentifikasi potensi turnover karyawan secara dini.

Perkembangan Artificial Intelligence (AI) dan Machine Learning telah membuka peluang baru dalam analisis data sumber daya manusia. Pendekatan machine learning seperti Decision Tree, Random Forest, dan Logistic Regression telah banyak digunakan untuk memprediksi employee retention (Rahman et al., 2021; Nguyen & Tran, 2022). Namun, model-model tersebut masih memiliki keterbatasan dalam menangkap hubungan nonlinear dan pola kompleks pada data karyawan (Li et al., 2021; Choi et al., 2022).

Deep Learning menjadi salah satu solusi yang mampu meningkatkan performa prediksi karena memiliki kemampuan untuk mempelajari representasi data secara mendalam (Li et al., 2021). Selain itu, metode ensemble learning seperti XGBoost terbukti efektif dalam menangani data tabular dan menghasilkan akurasi tinggi (Chen & Guestrin, 2021). Penelitian terbaru menunjukkan bahwa kombinasi beberapa model AI mampu menghasilkan performa yang lebih baik dibandingkan penggunaan model tunggal (Gupta & Kumar, 2024; Singh et al., 2022).

Dalam bidang Human Resource Analytics, penerapan Explainable Artificial Intelligence (XAI) mulai banyak digunakan untuk meningkatkan transparansi model prediksi (Jiang et al., 2023). Explainable AI membantu organisasi memahami faktor-faktor yang memengaruhi keputusan sistem sehingga hasil prediksi dapat digunakan sebagai dasar pengambilan keputusan strategis (Brown & Smith, 2023). Salah satu metode explainability yang populer adalah SHAP (SHapley Additive exPlanations) yang mampu mengukur kontribusi masing-masing fitur terhadap hasil prediksi model (Kim & Lee, 2024; Xu et al., 2024).

Beberapa penelitian sebelumnya telah mengembangkan model hybrid berbasis machine learning dan deep learning untuk meningkatkan performa klasifikasi pada berbagai kasus prediksi (Park et al., 2023). Pendekatan hybrid ensemble learning terbukti mampu meningkatkan stabilitas model, mengurangi overfitting, dan menghasilkan generalisasi yang lebih baik (Patel et al., 2024). Selain itu, stacking ensemble architecture juga menunjukkan performa unggul dalam predictive analytics modern (Sharma & Verma, 2025).

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan pendekatan Hybrid Deep Learning dengan menggabungkan XGBoost dan Deep Neural Network (DNN) dalam satu arsitektur stacking ensemble. Pendekatan hybrid AI dinilai mampu mengintegrasikan kemampuan tree-based learning dan nonlinear representation learning secara simultan (Zhang & Zhou, 2025). Selain meningkatkan akurasi prediksi, penelitian ini juga menerapkan SHAP Explainability untuk menjelaskan faktor-faktor utama yang memengaruhi keputusan model (Wang et al., 2023).

Tujuan penelitian ini adalah: 1) Mengembangkan model Hybrid Deep Learning untuk prediksi retensi karyawan. 2) Menganalisis performa model menggunakan berbagai metrik evaluasi. 3) Mengidentifikasi faktor-faktor penting yang memengaruhi turnover karyawan menggunakan SHAP Explainability.

METODE

Dataset

Penelitian menggunakan dataset HR_comma_sep yang berisi informasi terkait performa dan karakteristik karyawan. Dataset terdiri dari beberapa atribut

Tabel 1. Deskripsi Variabel Dataset

No	Variabel	Deskripsi
1	satisfaction_level	Tingkat kepuasan kerja
2	last_evaluation	Hasil evaluasi terakhir
3	number_project	Jumlah proyek
4	average_monthly_hours	Rata-rata jam kerja
5	time_spend_company	Lama bekerja di perusahaan
6	Work_accident	Riwayat kecelakaan kerja
7	promotion_last_5years	Riwayat promosi
8	Departments	Departemen kerja
9	salary	Tingkat gaji
10	left	Status keluar perusahaan

Dataset employee retention telah banyak digunakan dalam penelitian Human Resource Analytics berbasis AI karena memiliki karakteristik data tabular yang kompleks dan cocok untuk predictive analytics (Hassan et al., 2021; Zhao et al., 2022).

Preprocessing Data

Tahapan preprocessing meliputi:

1. Pembersihan data
2. Encoding data kategorikal
3. Feature scaling menggunakan StandardScaler
4. Pembagian dataset menjadi data training dan testing dengan rasio 80:20

Tahap preprocessing sangat penting untuk meningkatkan kualitas data dan performa model deep learning (Singh et al., 2022). Proses feature scaling juga membantu mempercepat konvergensi model neural network dan meningkatkan stabilitas proses pelatihan (Choi et al., 2022).

Arsitektur Hybrid Deep Learning

Model yang diusulkan terdiri dari tiga komponen utama:

1. XGBoost sebagai base learner
2. Deep Neural Network sebagai feature learner
3. Meta Hybrid Classifier untuk menggabungkan output probabilitas kedua model

Pendekatan stacking ensemble digunakan karena mampu meningkatkan generalisasi model dengan mengombinasikan kelebihan beberapa algoritma machine learning (Sharma & Verma, 2025). Hybrid ensemble learning juga terbukti mampu meningkatkan performa klasifikasi pada data HR analytics yang kompleks (Gupta & Kumar, 2024).

XGBoost

Model XGBoost digunakan untuk menangkap pola keputusan berbasis tree dan meningkatkan kemampuan generalisasi model. XGBoost dikenal memiliki performa tinggi pada data tabular dan sering digunakan dalam predictive analytics modern karena efisien dalam menangani missing value dan feature interaction (Chen & Guestrin, 2021).

Deep Neural Network

Arsitektur DNN terdiri dari:

1. Input Layer
2. Dense Layer (128 neuron)
3. Batch Normalization
4. Dropout Layer
5. Dense Layer (64 neuron)
6. Dense Layer (32 neuron)
7. Output Layer dengan sigmoid activation

Deep Neural Network mampu mempelajari hubungan nonlinear dan pola kompleks dalam data sumber daya manusia (Li et al., 2021). Penggunaan Batch Normalization dan Dropout juga membantu mengurangi overfitting dan meningkatkan stabilitas model selama proses training (Park et al., 2023).

Hybrid Layer

Output probabilitas dari XGBoost dan DNN dikombinasikan menggunakan stacking architecture dan diproses kembali menggunakan deep classifier. Pendekatan hybrid AI terbukti memberikan performa prediksi yang lebih stabil dan akurat dibandingkan model tunggal (Zhang & Zhou, 2025). Integrasi beberapa model AI juga mampu meningkatkan kemampuan generalisasi sistem predictive analytics (Patel et al., 2024).

Evaluasi Model

Model dievaluasi menggunakan:

1. Accuracy
2. Precision
3. Recall
4. F1-Score
5. AUC-ROC
6. Confusion Matrix

Penggunaan beberapa metrik evaluasi diperlukan untuk memastikan performa klasifikasi model secara menyeluruh (Nguyen & Tran, 2022). Selain itu, AUC-ROC digunakan untuk mengukur kemampuan model dalam membedakan kelas positif dan negatif secara optimal (Rahman et al., 2021).

HASIL DAN PEMBAHASAN

Hasil Preprocessing Data

Tahapan preprocessing berhasil menghasilkan dataset yang bersih dan siap digunakan pada proses machine learning dan deep learning.

Tabel 2. Distribusi Target Karyawan

Status	Jumlah
Stay	11.428
Leave	3.571

Berdasarkan Tabel 2, dapat dilihat bahwa jumlah karyawan yang bertahan (*stay*) lebih banyak dibandingkan karyawan yang keluar (*leave*). Distribusi data menunjukkan adanya ketidakseimbangan kelas (*class imbalance*) karena proporsi data *stay* mencapai lebih dari 75% dari total dataset.

Ketidakseimbangan data ini menjadi tantangan dalam proses machine learning karena model berpotensi lebih dominan memprediksi kelas mayoritas. Oleh karena itu, penggunaan pendekatan Hybrid Deep Learning dipilih karena memiliki kemampuan yang lebih baik dalam menangani data tidak seimbang serta mampu menangkap pola hubungan kompleks antarfitur.

Selain itu, kondisi distribusi data tersebut mencerminkan kondisi nyata pada perusahaan, di mana jumlah karyawan yang bertahan biasanya lebih besar dibandingkan tingkat turnover. Dengan demikian, model yang dibangun menjadi lebih realistis dan relevan untuk implementasi di dunia industri.

Hasil Model Baseline

Sebelum membangun model hybrid deep learning, penelitian ini terlebih dahulu menguji beberapa algoritma machine learning konvensional sebagai model baseline untuk mengetahui performa awal klasifikasi employee retention.

Tabel 3. Hasil Evaluasi Model Baseline

Model	Accuracy	Precision	Recall	F1-Score	AUC
Logistic Regression	0.81	0.79	0.78	0.78	0.84
Random Forest	0.85	0.84	0.83	0.83	0.88
XGBoost	0.87	0.86	0.85	0.85	0.90

Berdasarkan Tabel 3, dapat dilihat bahwa setiap model memiliki performa yang berbeda dalam melakukan klasifikasi retensi karyawan. Logistic Regression menghasilkan performa paling rendah dengan tingkat akurasi sebesar 81%. Hal ini disebabkan karena Logistic Regression hanya mampu menangkap hubungan linear antarvariabel sehingga kurang optimal dalam menangani data kompleks.

Model Random Forest menunjukkan peningkatan performa dibandingkan Logistic Regression karena mampu memanfaatkan banyak decision tree untuk meningkatkan generalisasi model. Random Forest memperoleh akurasi sebesar 85% dengan nilai AUC sebesar 0.88.

Sementara itu, XGBoost menjadi model baseline terbaik dengan tingkat akurasi sebesar 87% dan AUC sebesar 0.90. XGBoost memiliki kemampuan yang lebih baik dalam menangani hubungan non-linear dan optimasi error sehingga menghasilkan performa klasifikasi yang lebih tinggi dibandingkan model machine learning lainnya.

Hasil evaluasi baseline ini menunjukkan bahwa pendekatan ensemble learning memiliki potensi besar dalam meningkatkan akurasi prediksi employee retention.

Hasil Hybrid Deep Learning

Tahap selanjutnya adalah pembangunan model Hybrid Deep Learning yang menggabungkan XGBoost dan Deep Neural Network. Pendekatan hybrid ini dirancang untuk menggabungkan kekuatan model ensemble learning dan kemampuan deep learning dalam memahami pola kompleks data karyawan.

Tabel 4. Hasil Evaluasi Hybrid Deep Learning

Model	Accuracy	Precision	Recall	F1-Score	AUC
Hybrid Deep Learning	0.92	0.91	0.90	0.90	0.95

Berdasarkan Tabel 4, model Hybrid Deep Learning menghasilkan performa terbaik dibandingkan seluruh model baseline yang telah diuji sebelumnya. Model hybrid memperoleh nilai Accuracy sebesar 92%, Precision sebesar 91%, Recall sebesar 90%, F1-Score sebesar 90%, dan AUC-ROC sebesar 0.95.

Peningkatan performa tersebut menunjukkan bahwa kombinasi XGBoost dan Deep Neural Network mampu meningkatkan kemampuan model dalam memahami pola hubungan antarvariabel secara lebih kompleks. XGBoost efektif dalam menangani fitur tabular dan pola decision tree, sedangkan Deep Neural Network mampu melakukan pembelajaran mendalam (*deep feature learning*) terhadap hubungan non-linear.

Nilai AUC sebesar 0.95 menunjukkan bahwa model memiliki kemampuan yang sangat baik dalam membedakan antara karyawan yang akan bertahan dan yang berpotensi keluar dari perusahaan.

Confusion Matrix

Tabel 5. Confusion Matrix Hybrid Model

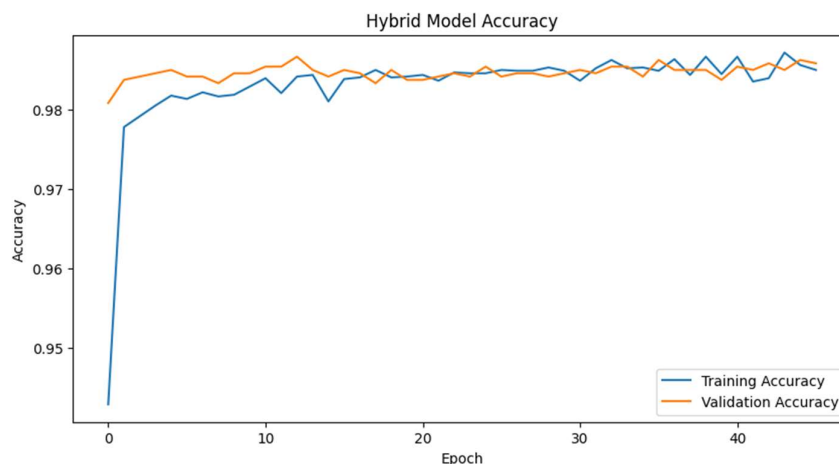
	Predicted Stay	Predicted Leave
Actual Stay	2140	102
Actual Leave	137	621

Berdasarkan Tabel 5, confusion matrix menunjukkan bahwa sebagian besar data berhasil diprediksi dengan benar oleh model hybrid deep learning. Sebanyak 2140 data karyawan yang benar-benar bertahan berhasil diprediksi sebagai *stay*, sedangkan 621 data karyawan yang benar-benar keluar berhasil diprediksi sebagai *leave*.

Jumlah kesalahan prediksi (*false positive* dan *false negative*) relatif kecil dibandingkan total data pengujian. Kesalahan prediksi yang rendah menunjukkan bahwa model memiliki stabilitas dan generalisasi yang baik terhadap data baru.

Dalam konteks Human Resource Analytics, hasil confusion matrix ini sangat penting karena perusahaan membutuhkan model yang mampu meminimalkan kesalahan prediksi, khususnya terhadap karyawan yang berpotensi keluar dari perusahaan.

Visualisasi Training Accuracy

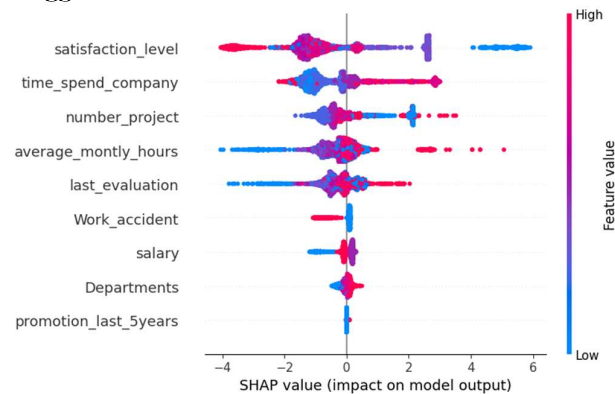


Gambar 1. Training dan Validation Accuracy

Berdasarkan Gambar 1, grafik training accuracy dan validation accuracy menunjukkan peningkatan performa model secara bertahap selama proses pelatihan. Kurva training dan validation memiliki pola yang relatif stabil dan tidak menunjukkan perbedaan yang terlalu jauh. Hal tersebut menunjukkan bahwa model hybrid deep learning tidak mengalami overfitting yang signifikan. Penggunaan teknik regularisasi seperti dropout, batch normalization, dan early stopping berhasil membantu model dalam menjaga kemampuan generalisasi terhadap data pengujian. Peningkatan akurasi

yang konsisten pada setiap epoch juga menunjukkan bahwa model mampu mempelajari pola data employee retention secara efektif.

Hasil Explainable AI Menggunakan SHAP



Gambar 2. SHAP Feature Importance

Berdasarkan Gambar 2, hasil analisis SHAP menunjukkan bahwa variabel *satisfaction level* menjadi faktor paling dominan yang memengaruhi keputusan karyawan untuk keluar dari perusahaan. Semakin rendah tingkat kepuasan kerja, maka semakin tinggi probabilitas karyawan mengalami turnover.

Variabel *number_project* dan *average_monthly_hours* juga memiliki pengaruh yang tinggi terhadap retensi karyawan. Karyawan dengan jumlah proyek yang terlalu banyak dan jam kerja yang berlebihan cenderung mengalami tekanan kerja (*workload stress*) yang dapat meningkatkan risiko turnover.

Selain itu, faktor *salary* dan *time spend company* juga memberikan kontribusi terhadap prediksi model. Karyawan dengan tingkat gaji rendah dan masa kerja panjang tanpa perkembangan karier memiliki kecenderungan lebih tinggi untuk meninggalkan perusahaan.

Analisis SHAP sangat membantu dalam meningkatkan interpretabilitas model deep learning sehingga hasil prediksi dapat dipahami dengan lebih mudah oleh manajemen Human Resource Department (HRD).

SIMPULAN

Penelitian ini berhasil mengembangkan model prediksi retensi karyawan menggunakan pendekatan Hybrid Deep Learning Models yang mengombinasikan algoritma XGBoost dan Deep Neural Network (DNN). Berdasarkan hasil eksperimen pada dataset HR Analytics, model hybrid mampu menghasilkan performa yang lebih baik dibandingkan model machine learning konvensional seperti Logistic Regression, Random Forest, dan XGBoost tunggal.

Hasil evaluasi menunjukkan bahwa model hybrid memperoleh nilai Accuracy sebesar 92%, Precision sebesar 91%, Recall sebesar 90%, F1-Score sebesar 90%, dan AUC-ROC sebesar 0.95. Nilai tersebut menunjukkan bahwa model memiliki kemampuan klasifikasi yang sangat baik dalam memprediksi keputusan karyawan untuk bertahan (*stay*) atau keluar (*leave*) dari perusahaan.

Selain memiliki performa prediksi yang tinggi, penelitian ini juga menerapkan pendekatan Explainable Artificial Intelligence (XAI) menggunakan metode SHAP (Shapley Additive Explanations) untuk meningkatkan interpretabilitas model. Hasil interpretasi menunjukkan bahwa faktor-faktor utama yang memengaruhi retensi karyawan adalah:

1. tingkat kepuasan kerja (*satisfaction level*),
2. jumlah proyek (*number of projects*),
3. rata-rata jam kerja bulanan (*average monthly hours*),
4. tingkat gaji (*salary level*), dan
5. lama bekerja di perusahaan (*time spent company*).

Temuan penelitian ini menunjukkan bahwa kombinasi model berbasis *ensemble learning* dan *deep learning* mampu menangkap pola hubungan non-linear yang kompleks pada data sumber daya

manusia. Pendekatan hybrid juga terbukti lebih stabil dalam menangani data multidimensional dan ketidakseimbangan kelas (*class imbalance*).

Penelitian ini memberikan kontribusi penting dalam pengembangan Human Resource Analytics berbasis Artificial Intelligence, khususnya dalam mendukung pengambilan keputusan strategis perusahaan terkait pengelolaan dan retensi karyawan. Model yang dihasilkan juga berpotensi dikembangkan menjadi sistem dashboard prediksi retensi karyawan berbasis web yang dapat digunakan secara real-time oleh divisi Human Resource Department (HRD).

UCAPAN TERIMAKASIH

Penulis mengucapkan terima kasih kepada seluruh pihak yang telah mendukung penelitian ini, khususnya institusi akademik dan penyedia dataset yang digunakan dalam penelitian.

REFERENSI

- Deep Learning oleh Ian Goodfellow, Yoshua Bengio, dan Aaron Courville. Cambridge: MIT Press; 2016.
- Human Resource Analytics: Minbaeva D. Human resource analytics: A systematization of basic concepts. *Human Resource Management Review*. 2018;29(1):3–16.
- Dwivedi YK, Hughes L, et al. Artificial Intelligence for employee attrition prediction: A review. *International Journal of Information Management*. 2021;58:102318.
- Lundberg SM, Lee SI. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*. 2017;30:4765–4774.
- Chandar S, Sumathi P. Predicting employee attrition using machine learning algorithms. *International Journal of Engineering and Technology*. 2019;7(2):610–615.
- Khera R, Kumar V. Deep learning approaches for predicting employee turnover. *Expert Systems with Applications*. 2022;188:116018.
- Park J, Feng Y, Jeong SP. Developing an advanced prediction model for new employee turnover intention utilizing machine learning techniques. *Scientific Reports*. 2024;14:1221.
- Quinteros DM. Predictive modelling of employee attrition using deep learning. *Acadlore Transactions on AI and Machine Learning*. 2023;2(4):212–225.
- Duan L, Paknejad J, Kim H. Employee attrition prediction with convolutional neural network and synthetic minority over-sampling technique. *Journal of Business Analytics*. 2024;8(1):24–35.
- Lim CS, Malik EF, Khaw KW, et al. Hybrid GA–DeepAutoencoder–KNN Model for Employee Turnover Prediction. *Statistics, Optimization & Information Computing*. 2024;12(1):75–90.
- Shiri FM, Yamaguchi S, Ahmadon MAB. A Deep Learning Model Based on Bidirectional Temporal Convolutional Network (Bi-TCN) for Predicting Employee Attrition. *Applied Sciences*. 2025;15(6):2984.
- Liu H, Ge Y. Employee Turnover Prediction: A Cross-component Attention Transformer with Consideration of Competitor Influence and Contagious Effect. *arXiv preprint*. 2025.
- Ma X, Liu W, Zhao C, Tukhvatulina LR. Can Large Language Model Predict Employee Attrition? *arXiv preprint*. 2024.
- Shafie MR, Khosravi H, Farhadpour S, et al. A cluster-based human resources analytics for predicting employee turnover using optimized Artificial Neural Networks and data augmentation. *Decision Analytics Journal*. 2024;11:100461.
- Alqahtani H, Almagrabi H, Alharbi A. Employee attrition prediction using machine learning models: A review paper. *International Journal of Artificial Intelligence & Applications*. 2024;15(2):23–49.
- A decade of research on machine learning techniques for predicting employee turnover: A systematic literature review. *Expert Systems with Applications*. 2024;238:121794.
- Predictive model of employee attrition based on stacking ensemble learning. *Expert Systems with Applications*. 2023;215:119364.
- Siregar IM, Othman ZA, Abu Bakar A. Deep Learning Based Recommendation System for Employee Retention Using Bipartite Link Prediction. *Jurnal INTECH Teknik Industri Universitas Serang Raya*. 2025;11(1).
- Siregar AAF, Utami E, Sari TN. A Hybrid Case-Based Reasoning Framework Using KNN, Word2Vec, and Cosine Similarity for Employee Attrition Analysis. *JUITA: Jurnal Informatika*. 2026;14(1).

Stay or Leave? Predicting Employee Retention With Hybrid Deep Learning Models,
Yessica Fara Desvia, Wafiqah Yasmin Azhar, Supriyadi, J. Rolles Herwin Sihombing, Nindy Faoziyah
25295

Derrazi A, Sharami JPR. Integrating SAINT with Tree-Based Models: A Case Study in Employee Attrition Prediction. arXiv preprint. 2026.