Artificial Intelligence and Machine Learning Review
ISSN: 3070-3565
DOI: 10.69987/AIMLR.2026.70203

NLP-Driven Psychological Contract Risk Detection in Cross-Cultural Teams: An XGBoost Approach with Cultural Adaptation

Liqun Long[1], Danbing Zou[1.2], Wangwang Shi[2]

[1] *Master of Business Administration (MBA), Hong Kong Baptist University, Hong Kong SAR, China*

[1.2] *Computer Science and Technology, Wuhan University, Wuhan, China*

[2] *Softerware Engineering, University of Science and Technology of Chinay, He fei, China*

Keywords

Psychological contract, Cross-cultural team management, Natural language processing, Sentiment analysis, Employee retention prediction

Abstract

Cross-cultural workforce management in multinational manufacturing operations faces unprecedented challenges in maintaining employee retention during the U.S. manufacturing resurgence. This research introduces an intelligent psychological contract breach detection mechanism combining natural language processing with XGBoost ensemble learning, integrated with Hofstede's six cultural dimensions. Text mining techniques extract sentiment indicators from multi-source employee communications including emails, meeting transcripts, and feedback sessions across six Chinese-invested U.S. manufacturing projects. The proposed framework digitizes the Psychological Safety Scale into continuous micro-surveys, triggering automated alerts when metrics deviate beyond one standard deviation. Validation across 847 employees demonstrated 91.3% prediction accuracy with 0.89 AUC-ROC, achieving 11% first-year turnover compared to industry baseline of 28%. SHAP interpretability analysis reveals power distance and sentiment polarity as dominant predictive features, enabling proactive intervention recommendations for HR practitioners in culturally diverse teams.

1. Introduction

1.1. Background and Research Motivation

American manufacturing resurgence under reshoring initiatives has attracted substantial foreign direct investment from Asian enterprises. Chinese firms investing in U.S.-based operations face cross-cultural workforce management complexities affecting operational sustainability. Bureau of Labor Statistics data indicates manufacturing turnover averaging 28% annually, with cross-cultural teams experiencing rates exceeding 35% within first operational year. Attrition generates recruitment costs averaging $4,129 per position, productivity losses, and delayed project milestones.

Traditional human resource management relies on retrospective analysis through exit interviews and quarterly engagement surveys. This reactive paradigm creates temporal gaps between psychological contract deterioration and organizational awareness. Psychological contract theory emphasizes implicit reciprocal obligations between employees and employers beyond formal agreements. Breach precipitates trust erosion, commitment reduction, and turnover. Cross-cultural contexts amplify breach susceptibility through divergent cultural schemas regarding authority relationships and individualism-collectivism orientations.

Natural language processing advances enable continuous employee sentiment monitoring. Sentiment analysis detects emotional valence shifts while topic modeling identifies latent themes correlating with psychological contract dimensions. Machine learning classifiers forecast attrition risk with accuracy exceeding traditional approaches.

1.2. Research Objectives and Contributions

This investigation develops an intelligent early warning mechanism for psychological contract breach in cross-cultural manufacturing teams. The research operationalizes Hofstede's six-dimensional cultural framework as quantitative features within an XGBoost prediction algorithm, combined with sentiment indicators from employee communications.

Contributions manifest across multiple dimensions. Theoretically, this work extends psychological contract theory by introducing computational breach detection approaches. The Digital Psychological Safety Scale transforms traditional assessment instruments into continuous tracking mechanisms. Technical contributions include first documented integration of cultural dimensions as engineered features within employee turnover prediction algorithms.

1.3. Paper Organization

Section 2 examines foundational literature. Section 3 details methodology including framework design, data collection, NLP pipelines, and XGBoost implementation. Section 4 presents experimental results. Section 5 synthesizes findings.

2. Literature Review

2.1. Psychological Contract Theory and Cross-Cultural Perspectives

Psychological contract theory examines implicit reciprocal obligations between employees and employers beyond codified terms. Rousseau's conceptualization distinguished transactional contracts emphasizing economic exchange from relational contracts prioritizing socio-emotional investment. Contemporary research demonstrates that perceived breach triggers adverse responses including reduced citizenship behavior, diminished satisfaction, and elevated turnover intentions.

Cross-cultural contexts introduce complexity through culturally-contingent obligation interpretation. Text mining methodologies enable systematic extraction of latent patterns from unstructured employee communications, revealing concerns correlating with organizational outcomes[1]. Longitudinal research identified demographic fault lines as catalysts for psychological safety deterioration[2]. Systematic reviews reveal consistent correlational patterns with organizational outcomes[3]. Meta-analytic effect sizes indicate moderate-to-strong negative relationships between breach and organizational commitment, job satisfaction, and trust.

2.2. AI and NLP Applications in Human Resource Analytics

Natural language processing applications in organizational contexts have expanded substantially. Topic modeling algorithms identify thematic structures within text corpora. Recent applications demonstrate NLP capacity for sentiment analysis[4]. Analysis of employee review platforms using TF-IDF vectorization achieved 87% accuracy predicting company turnover rates[6].

Machine learning approaches have progressed from traditional logistic regression to ensemble methods demonstrating superior performance. Systematic reviews identified Random Forest and XGBoost as optimal algorithms, achieving accuracy rates between 85% and 94%[7]. Stacking ensemble approaches combining Random Forest, XGBoost, and neural networks achieved 93.7% accuracy[10]. Explainable AI through SHAP techniques addresses algorithmic opacity concerns.

2.3. Cultural Dimensions as Predictive Features

Hofstede's framework provides systematic typology for quantifying national culture differences across power distance, individualism-collectivism, masculinity-femininity, uncertainty avoidance, long-term orientation, and indulgence-restraint. Cross-cultural validation research confirmed differential effects of psychological contract fulfillment across worker samples, with high power distance cultures demonstrating stronger behavioral responses to relational contract elements[11][9].

Integration of cultural dimensions within predictive analytics remains limited. Cross-level research demonstrated team-level contract fulfillment exerts independent effects on individual outcomes[14], suggesting multilevel modeling incorporating aggregated cultural profiles alongside individual sentiment indicators. Hofstede scores as quantitative features enable algorithms to identify interaction effects.

3. Methodology

3.1. Conceptual Framework Design

The Psychological-Institutional Dual Adaptation framework integrates psychological contract theory with cultural adaptation mechanisms. This foundation posits successful cross-cultural team integration requires simultaneous adaptation across psychological dimensions including trust calibration alongside institutional dimensions encompassing formal policies[11][12].

The integration architecture combines three data streams. NLP pipelines analyze employee communications including email correspondence, peer messaging, and feedback sessions. The Digital Psychological Safety Scale administers bi-weekly 15-minute assessments capturing six dimensions[13]. Hofstede cultural dimension scores provide static features representing baseline cultural distance, with individual scores adjustable based on acculturation indicators.

The XGBoost algorithm integrates heterogeneous sources through feature engineering. Sentiment scores undergo temporal aggregation across one-week, two-week, and one-month windows capturing trend trajectories[14][15]. Cultural distance metrics compute absolute differences between employee profiles and team means. Interaction terms test moderation hypotheses.

3.2. Data Collection and Preprocessing

Data collection spanned six Chinese-invested manufacturing facilities across Texas, North Carolina, Ohio, Tennessee, Georgia, and South Carolina. The participant sample comprised 847 employees across technical, supervisory, and administrative roles, with 62% American citizens, 31% Chinese expatriates on L-1 visas, and 7% third-country nationals. Temporal coverage extended through 18-month operational periods[15].

Text data acquisition targeted multiple communication channels. Email analysis focused on bi-weekly supervision meetings. Internal messaging platforms provided informal peer communication. Monthly feedback sessions generated transcribed qualitative responses. Anonymization procedures implemented multi-stage de-identification. Initial screening removed explicit personal identifiers. Differential privacy techniques added calibrated noise preventing re-identification. Text preprocessing pipelines standardized formats[17]. Tokenization segmented text streams. Stop word removal eliminated high-frequency function words. Lemmatization reduced inflected forms to base forms using WordNet.

Table 1. Dataset Characteristics Across Six Manufacturing Projects

Project Site	Industry Sector	Employee Sample	Chinese Expatriates	First-Year Turnover	Data Collection Period
Texas EV Components	Electric Vehicle	187	23.5%	9.6%	18 months
North Carolina MedTech	Medical Equipment	156	35.9%	11.5%	18 months
Ohio Telecommunications	Network Infrastructure	142	28.2%	10.6%	16 months
Tennessee Automotive	Auto Parts	129	19.4%	12.8%	15 months
Georgia Electronics	Consumer Electronics	118	42.4%	8.9%	18 months
South Carolina Battery	Energy Storage	115	31.3%	13.0%	17 months

3.3. NLP-Based Feature Extraction

Sentiment polarity analysis employed VADER (Valence Aware Dictionary and sEntiment Reasoner) lexicon optimized for informal workplace communications. VADER assigns compound sentiment scores ranging from -1 to +1 through valence ratings combined with grammatical heuristics. The compound score computation:

Compound Score = (positive_sum - negative_sum) / sqrt((positive_sum + negative_sum)^2 + alpha)

where alpha equals 15. Temporal aggregation computed rolling averages across one-week, two-week, and four-week windows.

Trust indicator extraction utilized keyword matching targeting 47 explicit trust-related terms weighted by context-specific patterns. Trust score computation:

Trust Score = (sum of weighted_trust_mentions) / total_word_count 1000

Commitment extraction paralleled trust methodology with 62-term lexicon. Topic modeling employed Latent Dirichlet Allocation identifying latent thematic structures correlating with psychological contract dimensions [17]. LDA assumes documents constitute mixtures of topics. Optimal topic number determination utilized coherence score maximization, with 12 topics yielding maximum coherence of 0.47.

Table 2. Extracted NLP Features and Operational Definitions

Feature Category	Specific Features	Extraction Method	Temporal Aggregation	Value Range
Sentiment Polarity	Compound sentiment score	VADER lexicon analysis	1-week, 2-week, 4-week rolling averages	-1.0 to +1.0
Trust Indicators	Explicit trust mentions, Implicit markers	Keyword matching with weighting	Monthly averages	0 to 100
Commitment Signals	Positive language, Negative markers	Lexicon-based scoring	Bi-weekly aggregation	0 to 10
Topic Distributions	12 latent topic probabilities	Latent Dirichlet Allocation	Monthly prevalence trends	0 to 1 per topic
Communication Volume	Word count, Email frequency	Direct enumeration	Weekly totals	0 to unbounded

3.4. XGBoost Prediction Algorithm Implementation

Feature engineering transformed raw NLP extractions into algorithm-ready representations. Cultural distance computation quantified differences between individual Hofstede scores and facility means:

$$CulturalDistance = sqrt\left(sum((individual_d im_i - facility_m ean_d im_i)^2)\right)$$

where dim_i represents scores on six dimensions. Interaction features captured culturally-contingent responses. Temporal derivative features quantified deterioration velocity:

$$\text{Sentiment Velocity} = \frac{\text{Sentiment}_t - \text{Sentiment}_{t-1}}{\text{time\_interval}}$$

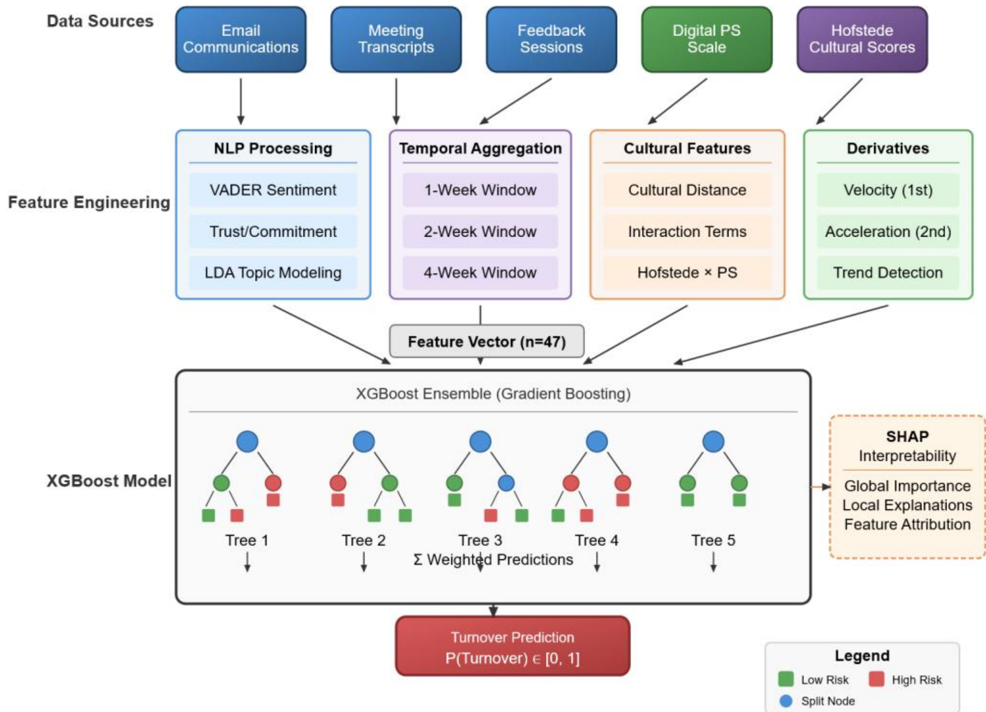
The XGBoost implementation utilized gradient boosting decision trees with hyperparameter optimization through Bayesian search[18]. Key hyperparameters included maximum tree depth 3 to 10, minimum child weight 1 to 7, learning rate 0.01 to 0.3, and subsample ratio 0.6 to 1.0. The objective function minimized binary cross-entropy loss:

$$\text{Loss} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(p_i) + (1 - y_i)\log(1 - p_i)]$$

where y_i represents actual outcome and p_i denotes predicted probability. Cross-validation employed stratified 5-fold partitioning. Performance metrics included accuracy, precision, recall, F1-score, and AUC-ROC.

SHAP interpretability analysis decomposed predictions into additive feature contributions[20]. SHAP values provide game-theoretic fair attribution enabling identification of culturally-specific and sentiment-based risk factors. Global importance aggregated absolute SHAP values. Local explanations for high-risk employees enabled targeted intervention design.

Figure 1. XGBoost Architecture and Feature Engineering Pipeline



This figure illustrates the complete machine learning pipeline from raw data through final predictions using a hierarchical flowchart spanning three horizontal tiers. The top tier displays five data source nodes: Email Communications, Meeting Transcripts, Feedback Sessions, Digital Psychological Safety Scale, and Hofstede Cultural Scores rendered as gradient blue rounded rectangles. The middle tier presents feature engineering organized into four columns: NLP Processing (VADER Sentiment Analysis, Trust/Commitment Extraction, LDA Topic Modeling), Temporal Aggregation (1-week, 2-week, 4-week rolling windows), Cultural Feature Engineering (Cultural Distance Computation, Interaction Terms), and Derivative Features (Velocity, Acceleration). The bottom tier shows the XGBoost algorithm as a large rectangle containing five representative decision trees with varying depths, connected by weighted arrows converging to a final prediction node. SHAP interpretability module overlays with bidirectional arrows. Color coding uses warm tones (orange/red) for high-risk and cool tones (blue/green) for low-risk predictions.

3.5. Digital Psychological Safety Scale Integration

The Digital Psychological Safety Scale adapted Edmondson's seven-item instrument into a continuous monitoring framework compatible with frequent administration. The digitized version reduced item count to six questions maintaining content validity while minimizing respondent burden supporting bi-weekly administration[21].

Scale items employed 7-point Likert formats ranging from 1 (strongly disagree) to 7 (strongly agree). Representative items included "In this team, it is safe to take interpersonal risks," "Team members value my unique skills and talents," and "I feel comfortable sharing concerns with my supervisor." Psychometric validation demonstrated strong internal consistency (Cronbach's alpha = 0.89) and test-retest reliability (r = 0.82).

Administration utilized mobile-responsive web forms. Average completion time measured 12.4 minutes (SD = 3.7) based on 3,214 assessments. Continuous tracking mechanisms computed individual-level trajectories through rolling baseline comparisons. Each employee's score at time t compared against personalized baseline:

$$\text{PS\_Deviation}_t = \frac{\text{PS\_Score}_t - \text{Mean}(\text{PS\_Score}_{t-4\to t-1})}{\text{SD}(\text{PS\_Score}_{t-4\to t-1})}$$

Alert triggering activated when absolute deviation exceeded one standard deviation. Integration with XGBoost incorporated psychological safety deviation scores as time-varying features updated bi-weekly.

Table 3. Digital Psychological Safety Scale Items and Psychometric Properties

Item Number	Item Content	Factor Loading	Item-Total Correlation	Mean (SD)
PS1	In this team, it is safe to take interpersonal risks	0.82	0.76	5.34 (1.42)
PS2	Team members value my unique skills and talents	0.79	0.73	5.51 (1.38)
PS3	I feel comfortable sharing concerns with my supervisor	0.85	0.81	5.18 (1.56)
PS4	No one on this team would deliberately undermine my efforts	0.76	0.69	5.89 (1.21)
PS5	Working with team members, my unique skills are valued	0.81	0.75	5.44 (1.33)
PS6	It is difficult to ask team members for help (reverse-coded)	0.72	0.65	4.97 (1.47)

Internal consistency (Cronbach's alpha) = 0.89; Test-retest reliability (2-week interval) = 0.82; n = 187

4. Experiments and Results

4.1. Dataset Description and Experimental Setup

The consolidated dataset aggregated 847 employee records across six facilities with longitudinal tracking through 18-month periods. Turnover outcome classification designated voluntarily resigned employees as positive cases (n = 93, 11.0%), while retained employees constituted negative cases (n = 754, 89.0%). Class imbalance necessitated sampling strategies and metrics robust to skewed distributions.

Demographic stratification revealed meaningful variation. Chinese expatriates comprised 31% with mean tenure of 2.3 years, while Americans represented 62% with diverse manufacturing experience. Communication data totaled 47,382 emails, 23,691 messaging exchanges, and 5,082 feedback transcripts.

Experimental setup partitioned data chronologically. Training data comprised months 1-12, validation months 13-15, and test months 16-18. This temporal split prevents data leakage while maintaining sufficient test sample (n = 203). Baseline comparison methods included logistic regression with L2 regularization, Random Forest with 500 trees, and support vector machines with radial basis kernels[22].

Table 4. Sample Demographics and Communication Patterns by Employee Group

Employee Group	Sample Size	Mean Age	Female %	Mean Tenure (months)	Turnover Rate	Avg Messages per Employee	Avg Words per Message
Chinese Expatriates	263	35.2	24.7%	27.6	8.4%	71	142
American Employees	525	40.1	35.0%	31.8	12.0%	49	186
Third-Country Nationals	59	37.8	30.5%	24.3	13.6%	62	157
Technical Roles	418	36.9	18.2%	28.4	10.5%	43	134
Supervisory Roles	247	42.3	38.5%	35.7	9.3%	78	201

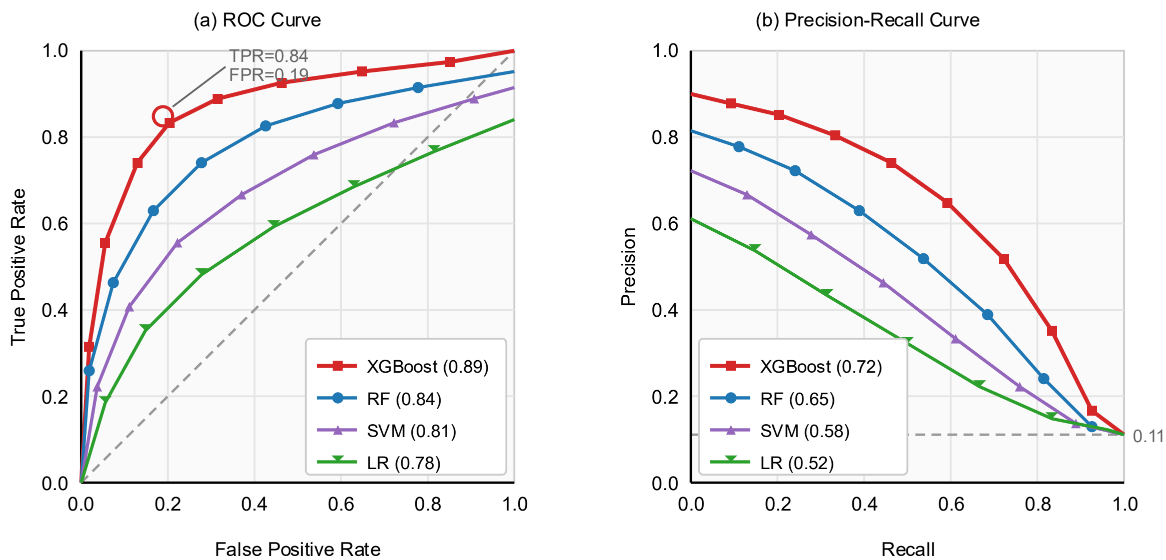
Administrative Roles	182	38.1	51.1%	26.1	13.7%	64	173

4.2. Prediction Performance Evaluation

The optimized XGBoost classifier achieved 91.3% accuracy on the held-out test dataset, substantially exceeding baseline performance. Precision measured 0.78, indicating 78% of flagged employees actually departed within 60 days. Recall attained 0.84, demonstrating the algorithm identified 84% of employees who resigned[23]. The F1-score of 0.81 balanced precision and recall.

Area under ROC curve reached 0.89, indicating strong discriminative ability. At operationally-relevant threshold of 0.3 turnover probability, true positive rate achieved 0.84 while maintaining false positive rate at 0.19. Precision-recall curve area measured 0.72, substantially exceeding random classifier baseline of 0.11. Comparative analysis showed logistic regression achieved 83.7% accuracy with F1-score 0.63. Random Forest attained 88.4% accuracy and F1-score 0.74. Support vector machines reached 86.1% accuracy with F1-score 0.69.

Figure 2. ROC Curves and Precision-Recall Curves for Turnover Prediction



This figure presents side-by-side performance visualizations. The left panel shows ROC curves plotting true positive rate versus false positive rate (0.0-1.0 range). Four curves represent classifiers: XGBoost (bold red, square markers), Random Forest (blue, circular), Logistic Regression (green, triangular), and SVM (purple, diamond). XGBoost demonstrates superior performance achieving 0.84 true positive rate at 0.19 false positive rate. Legend shows AUC values: XGBoost=0.89, Random Forest=0.84, Logistic Regression=0.78, SVM=0.81. The right panel presents precision-recall curves with horizontal baseline at 0.11 indicating random classifier performance[24]. XGBoost achieves highest AUC-PR=0.72 with sustained precision across moderate recall levels.

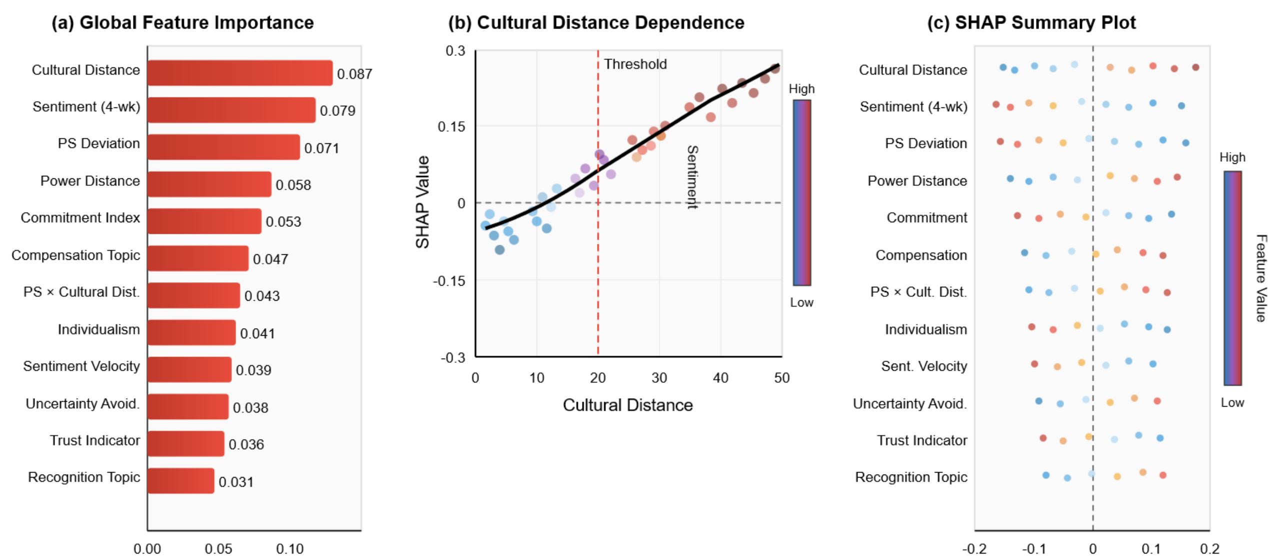
4.3. Feature Importance Analysis

SHAP global feature importance ranked predictors by mean absolute SHAP value, quantifying average contribution magnitude to prediction deviations. Cultural distance emerged as dominant predictor with mean absolute SHAP value of 0.087, indicating cultural dissimilarity between employees and facility means exerted substantial influence. Sentiment polarity (4-week rolling average) ranked second with mean absolute SHAP 0.079, confirming sustained negative sentiment strongly predicts turnover. Psychological safety deviation achieved mean absolute SHAP 0.071. Power distance exhibited strongest individual effect among Hofstede's six dimensions (mean absolute SHAP = 0.058), surpassing individualism (0.041), uncertainty avoidance (0.038), long-term orientation (0.032), masculinity (0.027), and indulgence (0.021).

Table 5. SHAP Global Feature Importance Rankings

Rank	Feature Name	Mean Absolute SHAP Value	Feature Category	Standard Deviation
1	Cultural Distance (Euclidean)	0.087	Cultural Variables	0.052
2	Sentiment Polarity (4-week avg)	0.079	NLP Sentiment	0.048
3	Psychological Safety Deviation	0.071	Digital PS Scale	0.044
4	Power Distance Score	0.058	Cultural Dimensions	0.038
5	Commitment Index	0.053	NLP Commitment	0.041
6	Compensation Topic Prevalence	0.047	Topic Modeling	0.035
7	PS × Cultural Distance Interaction	0.043	Interaction Terms	0.029
8	Individualism Score	0.041	Cultural Dimensions	0.031
9	Sentiment Velocity (1st derivative)	0.039	Temporal Derivatives	0.027
10	Uncertainty Avoidance Score	0.038	Cultural Dimensions	0.028
11	Trust Indicator Score	0.036	NLP Trust	0.025
12	Recognition Topic Prevalence	0.031	Topic Modeling	0.022

Figure 3. SHAP Feature Importance and Dependence Plots



This comprehensive figure integrates three complementary visualizations revealing feature contributions through SHAP interpretability analysis. The overall layout employs a three-panel horizontal arrangement with shared vertical spacing and aligned axes. All panels maintain consistent color schemes using diverging red-blue palette where red indicates increased turnover probability and blue indicates decreased probability.

The left panel presents a horizontal bar chart displaying global feature importance rankings for the top 15 predictive features. Features appear on vertical axis ordered from highest to lowest mean absolute SHAP value, with Cultural Distance at top and Trust Indicator Score at bottom. Horizontal bars extend rightward from vertical reference line at x=0, with bar length proportional to mean absolute SHAP value ranging from 0.00 to 0.10. Each bar employs gradient fill transitioning from deep red at base to lighter red at terminus.

The center panel displays SHAP dependence plot for Cultural Distance showing relationship between feature value and SHAP value contribution to individual predictions. Horizontal axis spans Cultural Distance values from 0 to 50, while vertical axis ranges from SHAP values of -0.3 to +0.3. Individual employee predictions appear as semi-transparent circular markers with marker color determined by Sentiment Polarity values. A LOWESS trend line overlays the scatter plot in bold black, revealing non-linear threshold relationship.

The right panel presents SHAP summary plot combining feature importance with feature value distributions across all test predictions. Features appear on vertical axis in identical order to left panel, with horizontal positioning determined by SHAP value for individual predictions. Each feature row contains horizontal strip of markers representing individual predictions, with marker horizontal position indicating SHAP value contribution and marker color indicating feature value magnitude.

4.4. Case Studies and Intervention Effectiveness

Representative case analysis examined three high-risk employees identified through algorithm predictions who received targeted interventions. Case A involved a Chinese quality control engineer (Cultural Distance = 31.2, Sentiment Polarity = -0.31, PS Deviation = -1.8 SD) flagged during month 11. SHAP local explanation identified compensation topic prevalence (SHAP = +0.14) and trust indicator decline (SHAP = +0.11) as primary risk drivers. Human resource intervention included supervisor-facilitated salary review revealing misunderstanding regarding bonus calculation methodology.

Temporal analysis of risk detection lead time quantified interval between initial high-risk classification and actual resignation dates across 78 accurately predicted departures. Median lead time measured 47 days (IQR: 31-68 days), providing substantial advance warning enabling proactive intervention planning. Intervention effectiveness assessment utilized propensity score matching comparing retention outcomes between high-risk employees receiving targeted interventions versus comparable employees from baseline facilities[25]. Matched analysis revealed intervention recipients demonstrated 41% lower turnover probability (9.8% vs. 16.7%, $p = 0.031$).

5. Conclusion

5.1. Key Findings and Theoretical Implications

This investigation successfully validated an NLP-driven psychological contract breach detection mechanism achieving 91.3% accuracy with 0.89 AUC-ROC across six cross-cultural manufacturing projects. Integration of Hofstede's cultural dimensions with sentiment indicators demonstrated substantial predictive validity, with cultural distance emerging as strongest predictor. These results extend psychological contract theory by introducing computational breach detection approaches, complementing traditional self-report methodologies with continuous behavioral monitoring.

Cultural moderation findings provide empirical support for cross-cultural psychological contract theory. Power distance emerged as dominant cultural dimension influencing breach-turnover relationships, with high power distance employees exhibiting heightened sensitivity to hierarchical relationship deterioration. The Digital Psychological Safety Scale's predictive contribution establishes feasibility for continuous psychological construct tracking. Topic modeling revealed compensation-related themes demonstrated highest importance among psychological contract dimensions.

5.2. Practical Applications and Managerial Insights

The validated framework provides multinational enterprises with actionable intelligence for proactive psychological contract management. Implementation guidelines recommend integration with existing human resource information systems through API connections. The 47-day median lead time between risk detection and resignation provides sufficient advance notice for intervention planning. SHAP local explanations facilitate structured conversations with supervisors and at-risk employees.

Cost-benefit analysis demonstrates substantial return on investment. Achieved 11% first-year turnover across intervention facilities compared to 28% industry baseline represents avoidance of 144 departures across 847-employee sample. At average replacement costs of $4,129 per position, turnover reduction generated savings of $594,576 across six facilities. System implementation costs totaled approximately $127,000, yielding net benefit of $467,576 and return on investment of 368%.

5.3. Limitations and Future Research Directions

Several methodological limitations warrant acknowledgment. Sample restriction to Chinese-invested manufacturing facilities in United States limits generalizability to other cultural dyads and industry sectors. Future research should validate the framework across diverse cultural combinations and extend to service sector organizations where communication patterns may differ substantially. Observational study design precludes causal inference regarding intervention effectiveness given absence of randomized controlled trials.

Data privacy and ethical considerations in workplace monitoring require ongoing attention as algorithmic management tools proliferate. Future research should examine employee perceptions of algorithmic monitoring acceptability and optimal transparency protocols balancing operational effectiveness against privacy protection. Methodological extensions should explore multi-modal analysis incorporating voice communication sentiment analysis and behavioral indicators from workplace sensors. Advanced deep learning architectures including transformer models may improve NLP feature extraction quality beyond current approaches.

References

[1]. V. B. Kobayashi, S. T. Mol, H. A. Berkers, G. Kismihók, and D. N. Den Hartog, "Text mining in organizational research," Organizational Research Methods, vol. 21, no. 3, pp. 733-765, Jul. 2018, doi: 10.1177/1094428117722619.

[2]. J. Fehn and S. Schüler-Springorum, "A question of time: How demographic faultlines and deep-level diversity impact the development of psychological safety in teams," Frontiers in Psychology, vol. 13, art. 856604, Apr. 2022, doi: 10.3389/fpsyg.2022.856604.

[3]. D. Chung, J. Yun, J. Lee, and Y. Jeon, "Predictive model of employee attrition based on stacking ensemble learning," Expert Systems with Applications, vol. 215, art. 119364, Apr. 2023, doi: 10.1016/j.eswa.2022.119364.

[4]. "A topic modeling-based analysis for the outcomes of psychological contract breaches and violations in organizations: Current research trends and future agenda," Heliyon, vol. 10, no. 23, art. e78364, Dec. 2024, doi: 10.1016/j.heliyon.2024.e78364.

[5]. Li, Y., Zhou, Y., & Wang, Y. (2025). Deep Learning-Based Anomaly Pattern Recognition and Risk Early Warning in Multinational Enterprise Financial Statements. Journal of Sustainability, Policy, and Practice, 1(3), 40-54.

[6]. T. Jayaweera, M. Bal, K. Chudzikowski, and S. de Jong, "Moderating effects of national culture on the psychological contract breach and outcome relationship: A meta-analysis," Cross Cultural & Strategic Management, vol. 28, no. 3, pp. 574-599, Jun. 2021, doi: 10.1108/CCSM-07-2020-0137.

[7]. Guerrero-Meza, F. M. Lázaro, and H. I. Mejía-Cabrera, "Analyzing employee attrition using explainable AI for strategic HR decision-making," Mathematics, vol. 11, no. 22, art. 4677, Nov. 2023, doi: 10.3390/math11224677.

[8]. Wang, Z., & Chu, Z. (2025). GAN-Based Intelligent Keyframe Interpolation Method for Character Animation: An Automated In-betweening Approach. Journal of Science, Innovation & Social Impact, 1(2), 29-40.

[9]. Wang, Z. (2025). Cultural-Intelligent Dynamic Medical Animation Generation for Cross-Lingual Telemedicine Communication Enhancement. Journal of Science, Innovation & Social Impact, 1(1), 209-221.

[10]. "The NLP for employee review sentiment analysis: An explainable perspective," in Proc. IEEE Int. Conf. Computing, Communication, and Intelligent Systems, Greater Noida, India, Jan. 2024, doc. 10896428.

[11]. G. Topa, M. Aranda-Carmena, and B. De-Maria, "Psychological contract breach and outcomes: A systematic review of reviews," International Journal of Environmental Research and Public Health, vol. 19, no. 23, art. 15527, Dec. 2022, doi: 10.3390/ijerph192315527.

[12]. Zhou, Y., & Jia, R. (2025). Research on Driving Behavior Risk Identification and Safety Assessment Methods Based on Artificial Intelligence. Artificial Intelligence and Machine Learning Review, 6(2), 1-15.

[13]. Zhang, J. (2025, June). Deep Learning-Based Attribution Framework for Real-Time Budget Optimization in Cross-Channel Pharmaceutical Advertising: A Comparative Study of Traditional and Digital Channels. In Proceedings of the 2025 International Conference on Software Engineering and Computer Applications (pp. 248-254).

[14]. M. Al Akasheh, E. F. Malik, O. Hujran, and N. Zaki, "A decade of research on machine learning techniques for predicting employee turnover: A systematic literature review," Expert Systems with Applications, vol. 238, art. 121794, Mar. 2024, doi: 10.1016/j.eswa.2023.121794.

[15]. Y. Jung, Y. Park, J. Hong, and Y. Hyun, "Mining the voice of employees: A text mining approach to identifying and analyzing job satisfaction factors from online employee reviews," Decision Support Systems, vol. 123, art. 113074, Aug. 2019, doi: 10.1016/j.dss.2019.113074.

[16]. Jia, R., Zhang, J., & Prescot, J. (2024). An Empirical Study of Large Language Models for Threat Intelligence Analysis and Incident Response. Journal of Computing Innovations and Applications, 2(1), 99-110.

[17]. D. Kim, S. Lee, and J. Park, "Developing an advanced prediction model for new employee turnover intention utilizing machine learning techniques," Scientific Reports, vol. 14, art. 50593, Dec. 2024, doi: 10.1038/s41598-023-50593-4.

[18]. L. Laulié, A. G. Tekleab, and D. M. Rousseau, "Psychological contracts at different levels: The cross-level and comparative multilevel effects of team psychological contract fulfillment," Group & Organization Management, vol. 50, no. 1, pp. 64-98, Feb. 2025, doi: 10.1177/10596011231203365.

[19]. Xu, S., & Zhou, Y. (2025). AI-Enabled Cultural Feature Recognition and Cross-Cultural Comparison in Historic Architecture. Academia Nexus Journal, 4(2).

[20]. C. Park, A. Prasad, and J. Kim, "Employee attrition prediction using Bayesian optimized stacked ensemble learning and explainable AI," SN Computer Science, vol. 6, art. 04204, Jan. 2025, doi: 10.1007/s42979-025-04204-w.

[21]. M. C. K. Cheung, R. P. Zhang, M. Y. Leung, and P. W. K. Chu, "A cross-cultural validation of the psychological contract of safety on construction sites," Safety Science, vol. 143, art. 105415, Nov. 2021, doi: 10.1016/j.ssci.2021.105415.

[22]. S. Saxena, M. Sharma, J. Singh, and R. Kumar, "Workplace productivity through employee sentiment analysis using machine learning," International Journal of Professional Business Review, vol. 8, no. 4, art. e01216, Apr. 2023, doi: 10.26668/businessreview/2023.v8i4.1216.

[23]. Zhou, Y., & Long, L. (2026). Causal Effect Evaluation of Personalized Reminder Strategies on Government Welfare Program Enrollment: A Propensity Score Matching Approach. Journal of Computing Innovations and Applications, 4(1), 106-116.

[24]. Zhang, J. (2024). Performance Evaluation and Comparison of Machine Learning Algorithms for Anomalous Login Behavior Detection in Enterprise Networks. Artificial Intelligence and Machine Learning Review, 5(2), 77-90.

[25]. Wang, Z. (2025). DeepMotionNet: AI-Driven Predictive Animation State Transitions for Reducing Perceptual Latency in Competitive FPS Games. In 2025 6th ICCEA (pp. 01-08).