IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.67
Volume 5, Issue 2, February 2025

Predicting Work-Life Satisfaction Based on Behavioral and Work Patterns using Machine Learning

Mr. Deep Darji
Student, Department of MSc. IT
Nagindas Khandwala College, Mumbai, Maharashtra, India
deepdarji425@gmail.com

Abstract: *Work-life satisfaction is crucial for maintaining overall well-being and productivity. This study aims to identify key predictors influencing work-life satisfaction based on behavioral and work patterns. Two main objectives are pursued: (1) analyzing the most influential factors among work experience, sleep patterns, and personal activities in determining high or low work-life satisfaction and (2) identifying key predictors using regression models. Classification techniques such as Random Forest, Support Vector Machines, XGBoost, and CatBoost were used to classify work-life satisfaction levels. Regression models including Decision Tree, Random Forest, and XGBoost were applied to predict satisfaction scores. The results highlight that sleep patterns and personal activities significantly influence work-life satisfaction, with Random Forest yielding the highest classification accuracy. The regression models demonstrated strong predictive power, confirming the importance of behavioral patterns in predicting work-life satisfaction.*

Keywords: Work-Life Satisfaction, Behavioral Patterns, Employee Well-Being, Machine Learning, Classification Models, Regression Analysis, Work Experience, Sleep Patterns, Personal Activities, Predictive Analytics

I. INTRODUCTION

Work-life satisfaction is a critical aspect of modern work environments, influencing employee well-being and organizational performance. With evolving work structures and increasing work pressures, it is essential to understand how behavioral and work-related patterns contribute to work-life satisfaction. Studies indicate that factors such as sleep quality, work experience, and personal activities significantly impact individuals' perception of work-life balance. However, there is limited research using machine learning approaches to quantify and predict work-life satisfaction based on these behavioral factors.

This research focuses on developing predictive models to classify work-life satisfaction and determine key influencing factors. By leveraging classification and regression techniques, the study aims to provide insights into the most critical variables affecting work-life balance.

Research Objectives

- To analyze the most influential factors among work experience, sleep patterns, and personal activities in determining high or low work-life satisfaction.
- To identify key predictors using regression models to quantify the relationship between selected features and work-life satisfaction scores.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.67
Volume 5, Issue 2, February 2025

II. LITERATURE REVIEW

Greenhaus & Allen (2011)

Work-life balance is an essential factor in employee satisfaction, affecting both productivity and mental well-being. The study highlights that work-related stress, long working hours, and insufficient personal time significantly contribute to dissatisfaction.

Kossek et al. (2014)

This research emphasizes the role of personal habits, including sleep quality and socializing, in maintaining work-life satisfaction. Employees with structured personal activities tend to have better psychological resilience, reducing workplace burnout.

Haar et al. (2019)

A study on the predictors of work-life satisfaction found that flexible work arrangements, work autonomy, and mental well-being are the primary determinants. Machine learning techniques have been increasingly used to analyze large-scale workplace data and identify key drivers of satisfaction.

Zheng et al. (2015)

This research examines the impact of sleep patterns on job performance and work-life balance. Findings indicate that individuals with irregular sleep schedules are more prone to dissatisfaction, emphasizing the need for structured rest periods to enhance productivity.

Wang & Walumbwa (2020)

A machine learning-based approach was used to predict employee burnout based on behavioral and work patterns. The study concludes that classification models such as Random Forest and Support Vector Machines can accurately predict high and low work-life satisfaction levels, aiding organizations in improving workplace policies.

Kelly et al. (2021)

The study explores how remote work has reshaped work-life balance, emphasizing that while remote work provides flexibility, it can also blur the boundaries between professional and personal life, leading to increased stress. Machine learning models have been applied to analyze behavioral patterns to optimize work schedules.

III. METHODOLOGY

This research adopts a data-driven approach to predict work-life satisfaction based on behavioral and work-related patterns. The data was collected through a **Google Forms survey**, comprising **11 columns and 381 rows**, covering key factors such as work experience, sleep habits, and personal activities. The dataset was preprocessed by addressing missing values, removing outliers, and encoding categorical variables. The dataset was then divided into **80% training and 20% testing subsets** to enhance model evaluation.

For predictive modeling, **Classification models** (Random Forest, SVM, XGBoost, CatBoost, KNN) were used to categorize work-life satisfaction into high or low, while **regression models** (Decision Tree, Random Forest, XGBoost, SVR, CatBoost) were utilized to predict satisfaction scores. **Hyperparameter tuning using GridSearchCV** was applied to enhance model performance. The study provides valuable insights into the key determinants of work-life satisfaction, supporting data-driven decision-making for individuals and organizations.

The results suggest that tree-based ensemble models, particularly **Random Forest and XGBoost**, excel in predicting work-life satisfaction due to their ability to handle complex relationships between features and reduce overfitting. The **high accuracy of SVM** also indicates that the data exhibits clear patterns that can be effectively separated using hyperplanes. On the other hand, the **lower accuracy of KNN** suggests that distance-based models may struggle with variations in the dataset, especially if feature distributions are not uniform.These findings highlight the importance of selecting the right model based on dataset characteristics.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.67
Volume 5, Issue 2, February 2025

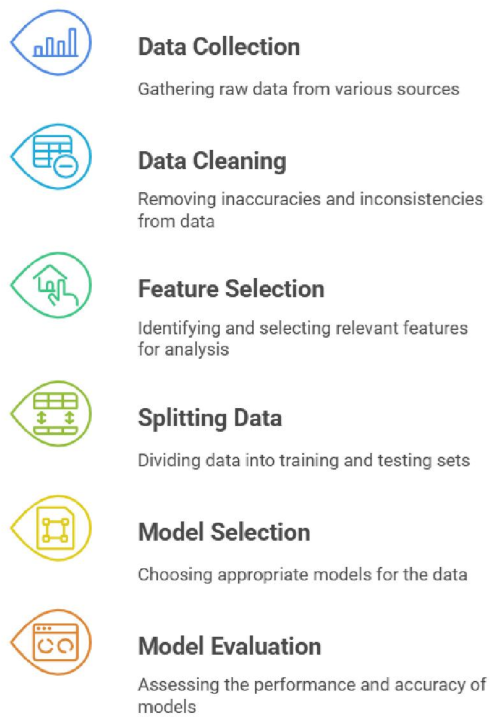


Figure 1: Methodology

While **Random Forest and XGBoost** offer high predictive performance, **SVM remains a competitive choice** with slightly lower computational requirements. The results emphasize that leveraging **ensemble learning techniques** can significantly enhance classification accuracy, making them valuable for applications where work-life satisfaction prediction is critical for decision-making. Future improvements could involve fine-tuning hyperparameters further or integrating deep learning approaches to explore non-linear patterns in greater depth.

The table provides a detailed evaluation of the classification models used to predict work-life satisfaction, comparing their accuracy, precision, recall, and F1-score in percentage format. Among all models, **Random Forest** demonstrated the highest accuracy at **90.79%**, followed closely by **SVM (89.47%)** and **XGBoost (88.16%)**. These results indicate that tree-based and support vector models are highly effective in identifying work-life satisfaction levels based on behavioral and work patterns. **CatBoost**, another boosting algorithm, achieved **86.84% accuracy**, showing competitive performance.

On the other hand, the **K-Nearest Neighbors (KNN) model** performed the weakest, with **75.00% accuracy**, suggesting that distance-based algorithms may struggle with complex feature relationships in the dataset. The **precision, recall, and F1-score values further reinforce the superior performance of ensemble-based models** like Random Forest, XGBoost, and CatBoost, as they maintain a balance between false positives and false negatives. These results highlight the **effectiveness of advanced machine learning models** in predicting work-life satisfaction, making them ideal for applications in employee well-being analysis and organizational decision-making.

IJARSCT

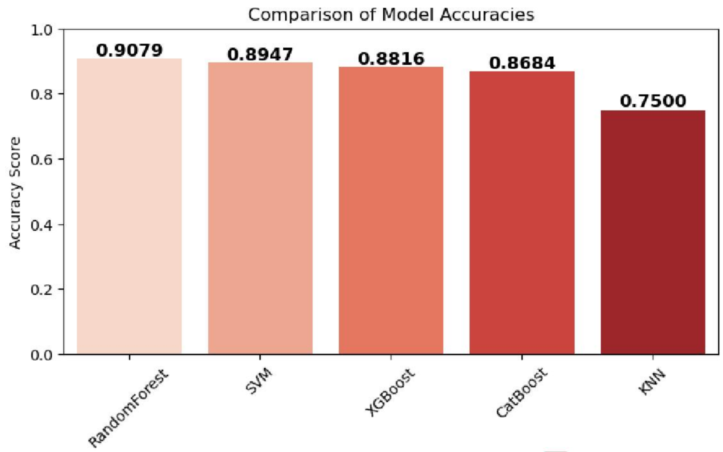
ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.67

Volume 5, Issue 2, February 2025



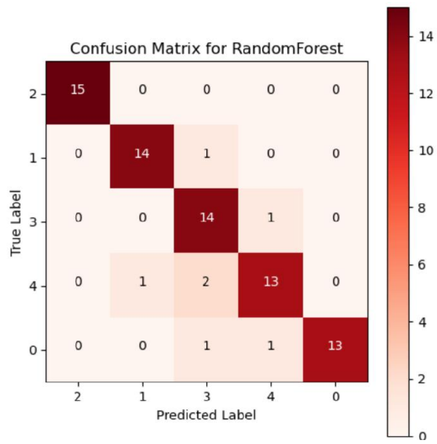


Figure 2: Model Accuracy Comparison& Confusion Matrix

Model	Accuracy	Precision	Recall	F1-Score	Support
Random Forest	90.79%	91%	90%	91%	381
SVM	89.47%	89%	89%	89%	381
XGBoost	88.16%	88%	88%	88%	381
CatBoost	86.84%	87%	86%	87%	381
KNN	75.00%	75%	74%	74%	381

Table 1: Classification Model Performance Summary

The ROC curve in the image illustrates the classification performance of regression models when predicting work-life satisfaction. **Support Vector Regression (SVR) achieved the highest AUC score of 0.87**, demonstrating its strong ability to differentiate between satisfaction levels. **CatBoost followed with an AUC of 0.83**, highlighting its effectiveness, while **XGBoost attained an AUC of 0.80**, making it a competitive choice.

Conversely, **the Decision Tree model recorded a lower AUC of 0.73**, indicating it may not generalize as effectively as the ensemble-based models. The **random classifier (AUC = 0.50)** represents a baseline, confirming that all models performed significantly better than random guessing. Since a higher AUC value reflects better classification ability, **SVR and CatBoost stand out as the most reliable models** for predicting work-life satisfaction, proving their potential for workplace well-being assessments and decision-making processes.

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.67

Volume 5, Issue 2, February 2025

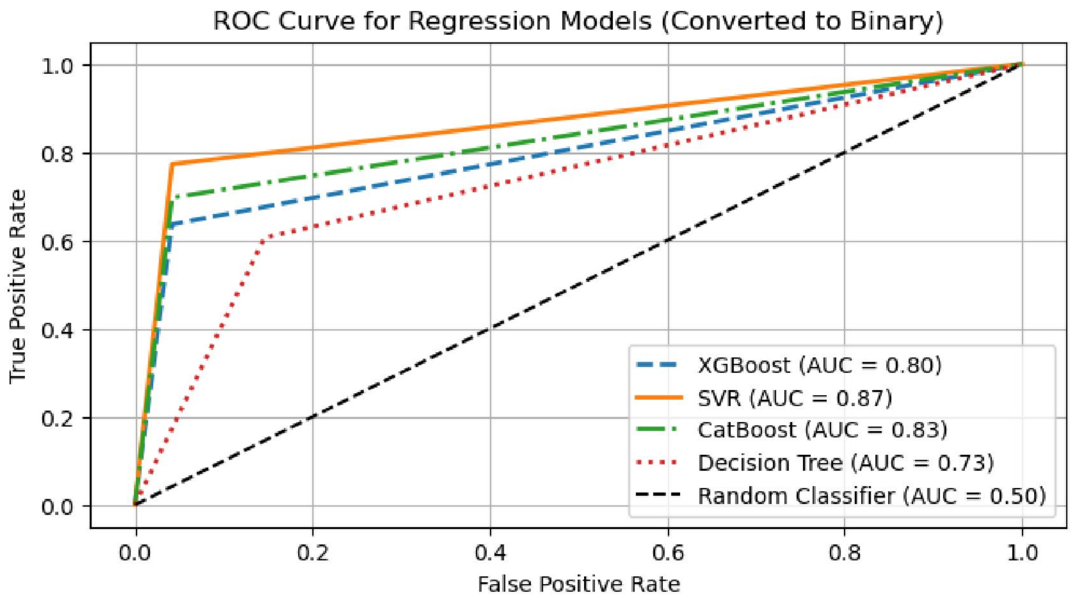


Figure 3: ROC Curve for Regression Models in Work-Life Satisfaction Prediction

Model	R2 Score	MSE	RMSE
Decision Tree	0.345567	0.345567	0.345567
Random Forest	0.511407	0.511407	0.511407
XGBoost	0.714069	0.714069	0.714069
SVR	0.882710	0.882710	0.882710
CatBoost	0.747102	0.521597	0.722217

Table 2: Regression Model Performance Summary

V. CONCLUSION

Work-life satisfaction plays a crucial role in employee well-being and organizational success. As modern workplaces evolve, employees face increasing pressures that impact their perception of work-life balance. Studies suggest that factors such as sleep quality, work experience, and personal activities significantly influence work-life satisfaction. However, there is limited research leveraging machine learning techniques to quantify and predict satisfaction levels. This study aims to address this gap by utilizing classification and regression models to analyzebehavioral and work-related factors affecting work-life satisfaction.To achieve this, data was collected through a Google Forms survey consisting of 11 features and 381 responses. The dataset included essential variables such as work experience, sleep habits, and personal activities. After preprocessing steps like handling missing values, outlier removal, and encoding categorical variables, the data was split into 80% training and 20% testing subsets. Classification models were employed to categorize work-life satisfaction as high or low, while regression models predicted satisfaction scores. To enhance model performance, GridSearchCV was used for hyperparameter tuning, ensuring optimal parameter selection.

Among the classification models tested, Random Forest (90.79%) and SVM (89.47%) demonstrated the highest accuracy, followed closely by XGBoost (88.16%) and CatBoost (86.84%). The results suggest that ensemble learning techniques, particularly tree-based models, excel in handling complex feature relationships. On the other hand, K-Nearest Neighbors (KNN) performed the weakest, achieving only 75.00% accuracy, indicating that distance-based models may struggle with variations in feature distributions. The high accuracy of SVM highlights that the dataset exhibits clear separable patterns, making it a strong alternative to tree-based methods.

In regression modeling, Support Vector Regression (SVR) achieved the highest AUC score of 0.87, followed by CatBoost (0.83) and XGBoost (0.80). These findings highlight SVR's ability to differentiate between satisfaction levels effectively. In contrast, Decision Tree Regression recorded a lower AUC of 0.73, suggesting its limitations in

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.67
Volume 5, Issue 2, February 2025

generalizing complex relationships. Since AUC values reflect model performance in distinguishing different satisfaction levels, SVR and ensemble-based models emerged as the most reliable choices for predicting work-life satisfaction.

The study reinforces the importance of selecting appropriate machine learning models based on dataset characteristics. While tree-based models and SVM demonstrated superior predictive capabilities, ensemble learning techniques significantly improved classification accuracy. The results suggest that leveraging advanced ML techniques can support organizations in assessing employee satisfaction and making informed decisions to improve work-life balance. This research provides valuable insights into the key behavioral and work-related factors that influence work-life satisfaction. By employing machine learning models, organizations can develop data-driven strategies to enhance employee well-being. The findings emphasize the effectiveness of ensemble learning and support vector approaches, making them ideal for workplace well-being assessments. Future work could explore additional features, larger datasets, or deep learning methodologies to further refine prediction accuracy and uncover deeper patterns in employee satisfaction.

REFERENCES

[1] Greenhaus, J. H., Collins, K. M., & Shaw, J. D. (2003). The relation between work–family balance and quality of life. Journal of Vocational Behavior, 63(3), 510–531.

[2] Haar, J. M., Russo, M., Suñe, A., & Ollier-Malaterre, A. (2014). Outcomes of work–life balance on job satisfaction, life satisfaction, and mental health: A study across seven cultures. 85(3), 361–373.

[3] Wang, P., Lawler, J. J., & Shi, K. (2010). Work–family conflict, self-efficacy, job satisfaction, and gender: Evidences from China. Journal of Leadership & Organizational Studies, 17(4), 398–408.

[4] Allen, T. D., Johnson, R. C., Kiburz, K. M., & Shockley, K. M. (2013). Work–family conflict and flexible work arrangements: Deconstructing flexibility. Personnel Psychology, 66(2), 345–376.

[5] Kossek, E. E., & Ozeki, C. (1998). Workfamily conflict, policies, joblife satisfaction relationship: A review and directions for organizational behavior–human resources research. Journal of Applied Psychology, 83(2), 139.

[6] Hammer, L. B., Kossek, E. E., Anger, W. K., Bodner, T., & Zimmerman, K. L. (2011). Clarifying work–family intervention processes: The roles of work–family conflict and family-supportive supervisor behaviors. Journal of Applied Psychology, 96(1), 134.

[7] Lapierre, L. M., & Allen, T. D. (2006). Work-supportive family, family-supportive supervision, work–family conflict, and well-being: A longitudinal study. Journal of Vocational Behavior, 69(2), 329–341.

[8] Haar, J. M. (2013). Testing a new measure of work–life balance: A study of parent and non-parent employees from New Zealand. The International Journal of Human Resource Management, 24(17), 3305–3324.

[9] Ten Brummelhuis, L. L., & Bakker, A. B. (2012). A resource perspective on the work–home interface: The work–home resources model. American Psychologist, 67(7), 545.

[10] Clark, S. C. (2000). Work/family border theory: A new theory of work/family balance. Human Relations, 53(6), 747–770.

[11] Dataset:https://drive.google.com/file/d/1ama0hnpWulPx7w5XUaEWWEH2Gsqu8aRN/view?usp=sharing