I.J. Modern Education and Computer Science, **2024, 6, 1-19**
Published Online on December 8, 2024 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijmecs.2024.06.01

Exploring Holistic Well-Being: EDA and a Comparative ML Approach to Work-Life Balance

Bala Dhandayuthapani V.
Department of IT, College of Computing and Information Sciences, University of Technology and Applied Sciences, Shinas campus, Oman
Email: bala.veerasamy@utas.edu.om, dhanssoft@gmail.com
ORCID iD: https://orcid.org/0000-0002-8310-0642

Received: 07 February, 2024; Revised: 30 March, 2024; Accepted: 16 May, 2024; Published: 08 December, 2024

Abstract: This research focuses on lifestyle and work-life balance, examining how different dynamics can influence holistic wellbeing, which focuses on the current literature to develop a framework that extends beyond standard well-being concepts. The work aims to enrich the discourse on holistic well-being by synthesizing existing knowledge and exploring new knowledge, providing valuable views for individuals and age groups involved in professionals striving to develop a more balanced and meaningful life. The study utilizes AI-driven methods and machine learning algorithms to analyse work-life indices to understand patterns and correlations that contribute or hinder work-life harmony. The findings highlight the importance of lifestyle choices, social connections, and personal fulfilment in achieving holistic wellness. The research provides evidence-based insights and practical recommendations to develop a healthy lifestyle and work-life balance. The study also examines the ethical implications of AI and highlights the need for a comprehensive approach to work-life balance. The study utilizes supervised learning algorithms and a comparative analysis of the accuracy scores of various algorithms, revealing significant differences in classification and more accuracy for the work-life balance score. The study aims to uncover insights beyond the typical contrasts and illuminate the interrelationship between numerous factors affecting an individual's well-being.

Index Terms: Exploratory Data Analysis, Holistic Well-being, Supervised Learning, Work-life balance, Work-life Harmony

1. Introduction

In modern workplaces, artificial intelligence (AI) has drastically changed the way human beings work, live, and take care of themselves. The conventional perception of work-life balance (WLB) has now evolved into work-life synchronization, where people aim to balance their practice and personal lives in a unified manner. The ethical considerations of openness, accountability, and automated decision-making have become critical in this environment. The fast-paced nature of AI-influenced work environments, along with the requirement for continual learning and adaptability, has the potential to have a wide-ranging impact on individual health. The ethical considerations surrounding AI have become central to discussions about responsible development and deployment. The ethical dimensions [1] of AI are crucial for navigating its impact on work and life. Employees working from home report higher job satisfaction and life satisfaction, but no significant difference in perceived stress. Employers should offer flexible work arrangements for better well-being. Holistic well-being (HWB), an encompassing concept that goes beyond the mere absence of illness, involves the pursuit of optimal health, happiness, and life satisfaction [2]. Achieving HWB is intricately linked to how individuals navigate their work obligations and personal lives.

The importance of balancing work-family and work-health responsibilities is emphasized [3] for overall well-being. Work-life harmony, a subtle perspective on work-life balance, reflects the dynamic integration of professional and personal spheres. Recognizing that individuals juggle diverse responsibilities, including work obligations and personal aspirations, this study aims to explore the cultivation of balance and harmony.

The significance [4] of WLB for subjective well-being (SWB) is highlighted, incorporating factors that moderate the relationship between WLB, job satisfaction, and SWB among Pakistani healthcare professionals, which suggests WLB significantly impacts SWB and job satisfaction, with intrinsic motivation moderating the relationship.

The relationship between AI and community well-being(CWB) was explored [5] in local governance. It highlights the potential of AI to improve societal issues and create sustainable, inclusive landscapes, emphasizes the importance of responsible and ethical AI approaches, including emotional AI, intellectual freedom, and healthcare. The significance of WLB [6] and working conditions in influencing self-reported health outcomes among adults in associations with age, and marital status. As the nature of work evolves and individuals face diverse challenges in insisting a balance between their practiced and individual lives, lifestyle choices become integral to the discussion. The impact of lifestyle on health is a complex interplay that extends beyond the workplace. Long working hours [7] destructively impact WLB, pleasure, stress levels, and social connections. Factors like job satisfaction, job security, and workplace characteristics influence this relationship. The relationship between WLB and mental and physical health among professionals was investigated [8], shedding light on the intricate dynamics at play. Mindfulness-based [9] interventions improve psychological detachment, WLB, and affective well-being, but not strain-based conflict or negative affect, which should explore long-term outcomes, boundary management strategies, and segmentation preferences.

The significance of WLB [6] and working conditions in influencing self-reported health outcomes among adults in association with age and marital status. As the nature of work evolves and individuals face diverse challenges in maintaining a balance between their professional and personal lives, lifestyle choices become integral to the discussion. The impact of lifestyle on health is a complex interplay that extends beyond the workplace. Long working hours [7] destructively influence WLB, pleasure, stress levels, and social connections. Factors like job satisfaction, job security, and workplace characteristics influence this relationship. The relationship between WLB and mental and physical health among professionals investigated [8], shedding light on the intricate dynamics at play. Mindfulness-based [9] interventions better psychological fairness of WLB, and affective well-being, but not strain-based conflict or negative affect, which should explore long-term outcomes, boundary management strategies, and segmentation preferences.

This study uses AI-driven methodologies to analyse work-life indices to understand patterns and correlations that contribute to or hinder work-life harmony and HWB. It aims to contribute evidence-based insights to the discourse on responsible AI implementation and its impact on individuals' lives. The research explores the relationship between work obligations, lifestyle choices, and overall well-being, raising concerns about ethical implications and the impact of AI on HWB.

Problem Statements: This research explores the relationship between lifestyle and WLB, aiming to provide comprehensive insights beyond conventional perspectives. It explores the intricate dynamics that shape HWB, aiming to illuminate how various aspects of an individual's life impact their well-being. The research aims to provide valuable implications for understanding and practical applications in enhancing life satisfaction and fulfilment.

RQ1: Low Correlated Features (Life-Style): How do lifestyle factors, including daily steps, body mass index (BMI) range, daily sleep hours, weekly meditation, fruits and veggies consumption, daily shouting, daily stress, places visited, core circle, sufficient income, and lost vacation, contribute to the perceived work-life balance score (WLBS) across different age groups?

RQ2: High Correlated Features (Work-Life): What is the impact of specific lifestyle and work-related aspects, such as time for passion, social network engagement, donation activities, supporting others, experiencing a state of flow, completing daily tasks (to-do completed), achieving personal goals, receiving personal awards, and having a clear life vision, on the perceived work-life balance score (WLBS) within various age groups?

RQ3: Work-life balance (WLB): What is the prevalent distribution of work-life balance score (WLBS) among the surveyed individuals, and how does this distribution reflect on the overall perception of work-life balance (WLB) within the studied population?

RQ4. Correlation Factors: What are low (Life-Style) and high (Work-Life) and selected correlated coefficient features?

RQ5: Work-life balance (WLB) and holistic well-being (HWB): What is the interplay between work-life balance (WLB) and holistic well-being (HWB)?

The above research questions lead to meaningful exploratory data analysis (EDA) from the work life dataset. The integration of AI into lifestyle and WLB can enhance efficiency and flexibility, resulting in healthier lifestyles. Lifestyle choices within work environments are essential for overall wellbeing. The research asserts that AI adoption is reasonable across industries, individual culture influences work-life harmony, and accurate, reliable data is utilized for analysis.

Research Objectives: The study uses a supervised learning algorithm to analyse the work-life dataset from Kaggle, focusing on WLB and well-being. It examines the ethical implications of AI use and the relationship between lifestyle and work obligations in AI-influenced settings. The research also explores the role of lifestyle choices in HWB in AI-enriched work environments. The findings will provide evidence-based insights for informed decision-making and practical recommendations for cultivating a balanced, health-conscious lifestyle and work-life ecosystem in the context of AI advancements.

Significances: The study investigates the connection between AI, lifestyle, WLB, and HWB. It provides insights into ethical considerations, work obligations, lifestyle choices, and health outcomes in AI-driven environments. The findings can inform individuals about the impact on personal well-being, promote healthy lifestyles, and increase productivity. It empowers employees to make informed choices and navigate work obligations, enhancing overall well-being. The research contributes to long-term social implications, balancing technological advances with human welfare.

Scope and limitations: The study aims to analyse lifestyle and WLB using AI-driven integration using machine learning (ML) and supervised learning algorithms. It utilizes EDA to examine the impact of AI on work-life harmony, ethical considerations, lifestyle choices, and well-being. The research will use cross-sectional and longitudinal analyses, considering diverse demographic factors and advanced data-driven methodologies. However, the research utilized the already available Kaggle dataset, but not real-time data gathered. And the findings may not be generalizable to all industries and organizational contexts due to the varying nature of AI integration, subjective ethical considerations, self-reported data, and external factors like economic conditions and societal changes.

2. Literature Review

Contemporary research on AI, lifestyle, WLB, and HWB has focused on the challenges and opportunities presented by technological advancements. This area encompasses theoretical underpinnings of personal WLB and practical implications of AI integration in the workplace. The study of HWB focuses on lifestyle choices, social connections, and personal fulfilment, examining WLB, mindfulness, structural working conditions, and interventions. The research reveals the subtle relationships within these dimensions, contributing to individuals' overall sense of well-being. A synthesis of key studies provides a comprehensive foundation for understanding the dynamics at play in contemporary work environments.

WLB and Well-being: WLB is a complex issue that involves the balance between professional and personal spheres, as well as the interplay between work-family and health considerations. Research highlights the importance of balancing these dimensions for overall health [3]. A comparative analysis of WLB [6] among European working adults, revealing that it is influenced by individual choices and societal structures. The relationship explored [7] between working time and well-being in Abu Dhabi, highlighting the structural effects of working time. The focus [8] is on specialists, managers, and entrepreneurs in Warsaw, examining the mental and physical health aspects of WLB. They emphasize the importance of achieving equilibrium between professional responsibilities and personal life to positively impact SWB. The role of WLB in influencing mental and physical health has been extensively studied.

Social Considerations in AI Integration: The intricate relationship between work and life has become a focal point of scholarly inquiry, with researchers delving into the subtle dynamics of WLB. A theoretical foundation, emphasizing the distinction between WLB [1] and the emerging concept of work-life merger. The rationale for this research underscores the evolving nature of work dynamics, prompting a critical exploration of how individuals navigate the boundary between professional and personal realms. Building upon this theoretical groundwork, they contribute empirical insights [4] by introducing a moderated mediation model to understand the balance between work and life and its impact on SWB. Their study recognizes the complex interplay of factors influencing well-being and posits that this relationship is contingent on certain moderating and mediating variables. By investigating SWB, they shed light on the psychological dimensions of WLB, further emphasizing its relevance in the contemporary professional landscape.

Impact of AI on Work-Life Harmony: Work-life dynamics have become a significant focus in contemporary work environments [1], distinguishing between WLB and the concept of work-life merger. A moderated mediation model (MMM) [4] to better understand the delicate balance between work and life and its effects on SWB. This model acknowledges that the relationship between work and life is not uniform across individuals and may be influenced by moderating and mediating factors. Further explored [7] the structural effects of working time on well-being, presenting evidence from Abu Dhabi. Their study introduces the notion that a balanced approach positively influences SWB. This research contributes to understanding how work obligations influence holistic health in diverse cultural and organizational contexts.

Mindfulness and Well-being: The study on mindfulness interventions [9] for WLB provides valuable insights into the benefits of cultivating a mindful approach in professional settings. The research examines the impact of segmentation preference on changes in detachment, well-being, and WLB. The study highlights the role of mindfulness in influencing individuals' mental compartmentalization of work and non-work aspects of life. The study adds a unique dimension to the literature on WLB by emphasizing the potential efficacy of mindfulness interventions in fostering a healthier balance between professional and personal spheres. The relationship between WLB and mental and physical health among professionals, emphasizing the need to consider individual lifestyle choices in the broader context of well-being.

Mindfulness and Work-Life Balance: Research has demonstrated that mindfulness therapies can improve WLB, with research showing the influence of segmentation preferences and structural factors of working hours on well-being. It identified [7] an association between working hours and overall well-being in Abu Dhabi. It investigated the effect of workplace participation [10] on accountants' WLB, demonstrating the complex relationships between workplace involvement and obtaining a balanced work-life. Individual successes and acts of charity have also been connected to well-being, with personal accolades and accomplishments increasing life happiness. It performed a comparative examination of gender and welfare state regimes [6] among working individuals in Europe, demonstrating differences in self-reported health across gender and welfare state regimes.

Employee Wellness Programs: The emphasizes the importance of employee wellness programs [11] for improved performance and overall well-being. These programs address not only individual well-being but also contribute to heightened performance outcomes. The study contributes to the growing literature on the strategic importance of HWB in the workplace, offering valuable insights into designing and implementing effective wellness programs that go beyond individual health benefits. It calls for ongoing research to refine and optimize the impact of these programs in the ever-changing landscape of modern workplaces.

Workplace and community well-being (CWB): The growing influence of AI on society requires strategic social design [12] for leveraging AI for CWB. It evaluates well-being within the current paradigm, highlighting its subjective, social, and psychological dimensions and its implications for economic growth. Workplace well-being is highlighted by its impact on employee health, productivity, and organizational success [13]. It highlights psychosocial relationships, job satisfaction, and social support as crucial factors and emphasizes the importance of a supportive work environment. WLB and social support examined [14] burnout among 73 social workers in two Indian cities, finding high rates and low social support. It suggests addressing burnout and compassion fatigue in social work curricula for improved mental health. The impact of WLB on employee engagement [15], incorporating motivation as a moderation variable, used a quantitative approach, using primary data from a questionnaire, and analysing it using smart PLS 4.0 software. The intersections of AI and CWB [16] aims to inform researchers, academics, practitioners, policymakers, and activists about the implications of AI on CWB. It addresses the growing reliance on AI and its impact on communities and ecosystems.

WLB and Job Satisfaction: Organizations should prioritize WLB initiatives to enhance job satisfaction and consider employees' needs at different career stages, as older employees reported higher WLB [17]. WLB [18] positively impacts employee performance in small and medium-sized enterprises (SMEs), with supervisor behaviours that are family-supportive enhancing these effects. WLB [19] and job satisfaction among higher education teachers, highlighting personal, organizational, and familial influences, suggests strategies, and emphasizes collaboration among researchers, policymakers, and teachers. The impact of WLB [20] on employee performance and job satisfaction, finding that flexible work arrangements, organizational support, and individual coping strategies positively influence it. The demographic factors [21] impact on WLB and job satisfaction among medical and dental professionals in Karachi, Pakistan, highlighting research gaps and suggesting that addressing these issues can enhance individual satisfaction and organizational growth.

The link between WLB, flexibility, psychological construct, and organizational culture on employee well-being highlights the need for a holistic approach [22]. A personalized employee work scheduling framework [23] uses metaheuristic algorithms to generate shift schedules based on employees' preferences. While swarm intelligence outperforms genetic algorithms, the research gap is in developing and evaluating this framework. Future research should explore different priorities. The relationship between social support, psychological well-being, and WLB in healthcare workers found that WLB is destructively associated to social support and positively connected to psychological well-being [24].

WLB Policies and Practices: Work-life balance [25] and its impact on employees' well-being, using a qualitative exploratory case study design aimed to improve conditions at home and work, especially post-COVID-19. The function of WLB [26] in employee loyalty, job satisfaction, and retention. However, research gaps exist on long-term effects, organizational culture, leadership support, and effectiveness evaluation, and addressing these can improve policy implementation. There is a strong correlation between WLB [27] and improved job functioning, highlighting the essential for an open insight of workplace culture and the environment's influence on WLB and productivity.

The literature brings enough study to understand WLB, which is a complex issue involving professional, personal spheres, work-family, job satisfaction, and work-health considerations. A comprehensive approach is essential for overall health and well-being. Studies have explored gender dimensions, the relationship between working life and well-being, and the integration of AI. Lifestyle choices have been found to contribute to well-being, and mindfulness interventions have been found to promote WLB.

3. Methodology

The study employs an exploratory data analysis (EDA) approach to investigate the relationship between work-life harmony, AI integration, and HWB. The data collection method is based on Kaggle datasets [28], which provide a broad spectrum of information on work-life dynamics, AI utilization metrics, and HWB indicators. The dataset contains 15,977 survey responses, each providing information on 24 qualities that characterize how we spend our lives. The data is categorized into age groups with each attribute accompanied by a question. The dataset includes the number of fruits or vegetables eaten daily, daily stress levels, visited places, core circle, supporting others, social network, achievement, donation, BMI range, to-do list completion, flow, daily steps, live vision, sleep hours, lost vacation days, daily shouting, sufficient income, personal awards, time for passion, and weekly meditation. The research aims to investigate the relationships between lifestyle choices and general well-being using the Kaggle dataset [28]. The research aims to identify correlations and patterns in various lifestyle factors such as nutrition, physical activity, sleep, and stress levels, providing insights into how these elements affect a person's well-being. The findings will also provide practical tips for people to improve their overall well-being through lifestyle changes.

ML workflow involves several essential phases to convert raw data into a prediction model. It starts with data collection, which involves acquiring relevant datasets from Kaggle sources [28]. The data preparation step ensures data acceptability for analysis. Feature scaling standardizes numerical characteristics, while feature selection and extraction select and maintain the most informative features. The model-building process involves selecting an appropriate method, training it on a predetermined subset, and fine-tuning it for peak performance. The model is then executed by applying it to fresh data for predictions, and cross-validation evaluates model resilience across different data subsets, as shown in Figure 1.

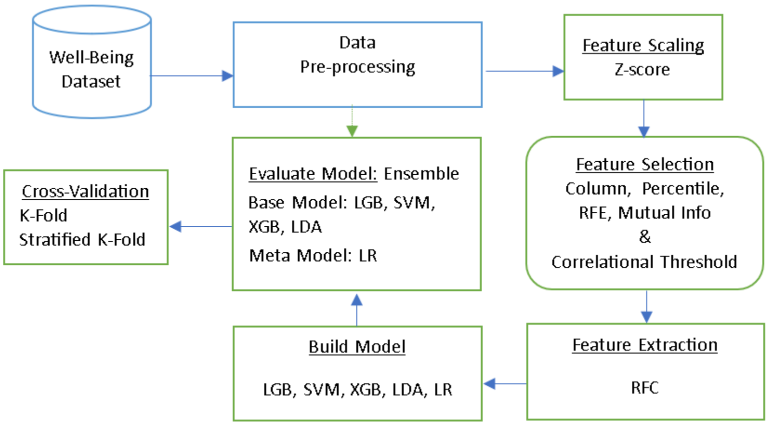

Fig. 1. Machine Learning Workflow Diagram

3.1 Data Pre-processing

Before starting the data exploration process, a data cleaning process is needed to work with this dataset with 15972 rows and 24 columns. It manages the datatypes, addresses inconsistencies, and manages duplicate entries and missing values. After investigation of the dataset, the process begins with the removal of the timestamp column, which is not necessary in this research and standardizing all the column names to lowercase for consistency of access during the analysis process. All rows with null values columns are removed for data completeness. The daily_stress and work_life_balance_score columns converted to numerical values. A row at index 10005 is deleted to maintain data integrity, which was holding a date value in the daily_stress column. The dataset, which contained 683 duplicate values, was removed to ensure data consistency and the trustworthiness of future studies. To facilitate further analysis, categorical variables such as age and gender are mapped to numerical values. 'Less than 20', '21 to 35', '36 to 50', '51 or more' are categorical age groups mapped with numerical values 0, 1, 2, and 3, respectively. 'Male' and 'Female' are categorical gender labels mapped to numerical values 0 and 1. The work_life_balance_score is then transformed into categorical bins based on predefined conditions, providing a more meaningful representation shown in Table 1. Finally, it ensures the reliability and readiness of the dataset for subsequent analytical tasks, which has 15288 rows with 23 columns.

Table 1. Mapping of Work-Life Balance Score (WLBS) with its Ranges

WLBS	WLBS Category	WLBS Ranges
0	Poor	<= 565
1	Average	> 565 & <= 650
2	Good	> 650 & <= 735
3	Very Good	> 735

Table 1 provided mapping outlines a clear categorization of the WLBS into distinct categories based on specified ranges. A WLBS value of 0 is categorized as 'Poor' and corresponds to scores less than or equal to 565. The category 'Average' is assigned to WLBS values of 1, falling within the range greater than 565 and less than or equal to 650. A WLBS value of 2 is labelled as 'Good' and encompasses scores greater than 650 but less than or equal to 735. Lastly, a WLBS value of 3 is classified as 'Very Good' for scores exceeding 735. This systematic mapping provides clarity and meaningful interpretation of WLBS values, allowing for a subtle understanding of WLB across different score ranges. Such categorization aids in simplifying the representation of WLBS data, making it more accessible and insightful for analytical purposes.

3.2 Feature Scaling

After the pre-processing steps are completed, the next step in the ML process is to feature scaling. A z-score or standard score compares a value to the mean of a set of values, determined in terms of standard deviations from the mean. The Python Pandas package and scikit-learn StandardScaler are utilized to standardize the dataset, which is stored in the data frame 'df'. The StandardScaler is used to determine the numerical features in the dataset, resulting in a mean of 0 and a standard deviation of 1. The standardized dataset, df_standardized, is then transformed into a data frame. Finally, a message is printed to indicate that the z-score transformation has been completed and that the standardized data frame is available for further analysis or modelling.

3.3 Feature Selection

Feature selection occurs after feature scaling is completed in the ML process. The feature selection processes used with various feature selection methods shown in Table 2. The SelectFromModel method chose WLBS as the relevant feature for predicting or classifying outcomes. The column-based approach selected the features places visited, supporting others, achievement, to-do completed, and WLBS based on their significance in the dataset. The percentile method also identified the same set of features as important for modelling: places visited, supporting others, achievement, to-do completed, and WLBS. The recursive feature elimination (RFE) yielded related results, identifying the same set of features as crucial. Lastly, the mutual information technique highlighted supporting others, achievement, to-do completed, time for passion, and WLBS as features with strong mutual information content. These selected features are likely to play a vital role in understanding and predicting WLBS, providing valuable insights into the factors influencing this crucial aspect of well-being.

Table 2. Feature Selection Methods and Selected Features

Feature Selection Method	Selected Features	Total Features
SelectFromModel	['work_life_balance_score']	1
Column-based	['places_visited', 'supporting_others', 'achievement', 'todo_completed', 'work_life_balance_score']	5
Percentile	['places_visited', 'supporting_others', 'achievement', 'todo_completed', 'work_life_balance_score']	5
Recursive Feature Elimination (RFE)	['places_visited', 'supporting_others', 'achievement', 'todo_completed', 'work_life_balance_score']	5
Mutual Information	['supporting_others', 'achievement', 'todo_completed', 'time_for_passion', 'work_life_balance_score']	5

The unique features in a dataset and the corresponding number of occurrences for each feature. Among these features, places visited appear three times, indicating their presence in three instances or occurrences within the dataset. Time for passion is a unique feature occurring once in the dataset. The WLBS is the most frequently occurring feature, appearing five times. Similarly, achievement, supporting others, and to-do completed each occur four times in the dataset. This distribution of unique features and their occurrences provides valuable insights into the dataset composition, shedding light on the varying importance and prevalence of specific features. Understanding the frequency of these features contributes to a comprehensive analysis, aiding researchers, or data analysts in making informed decisions regarding the significance of each feature in the dataset, as shown in Table 3 and the visual representation shown in Figure 2.

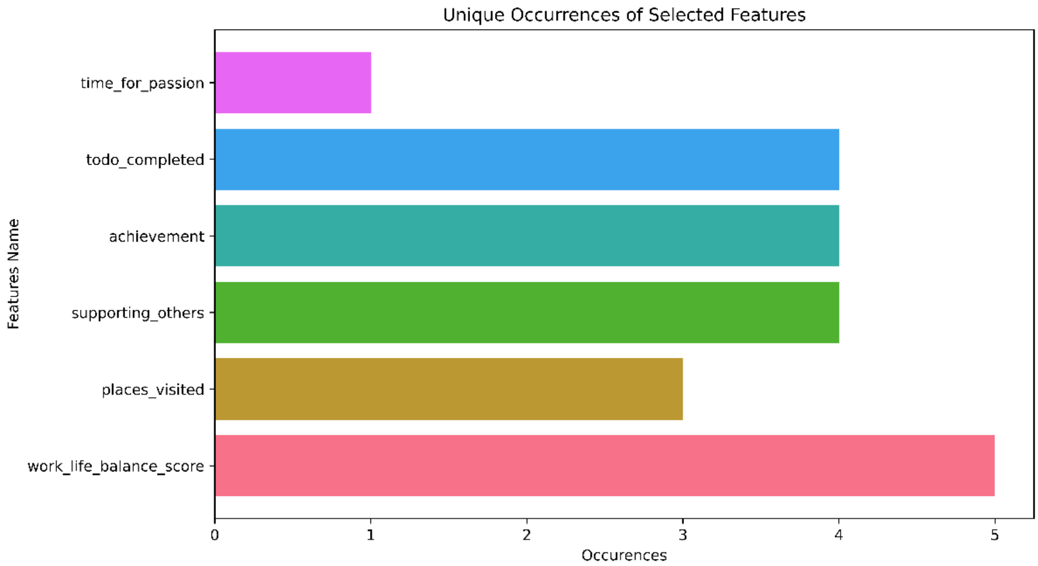

Fig. 2. Unique Occurrences of Selected Features

Table 3. Unique Occurrences of Selected Features

Unique Features	Number of Occurrences
time_for_passion	1
todo_completed	4
achievement	4
supporting_others	4
places_visited	3
work_life_balance_score	5

The correlation threshold value with 0.30 filtered the selection of highly correlated features within a dataset. The features of social network, personal awards, time for passion, supporting others, WLBS, achievement, donation, to-do completed, flow, and live vision exhibit notable correlation patterns. Specifically, the correlation matrix indicates a high degree of interdependence among these features, with column 10 and filtered column 13 showcasing strong correlations. Consequently, these features are considered highly correlated, suggesting that changes in one feature may be associated with changes in others within this subset. The remaining features are lower correlations with each other or do not exhibit strong correlation patterns with the features in the highly correlated subset. This subset includes features such as fruits veggies, daily stress, places visited, core circle, BMI range, daily steps, sleep hours, lost vacation, daily shouting, sufficient income, weekly meditation, age, and gender. Analysing features with varying correlation levels is crucial for gaining a comprehensive understanding of the dataset, informing subsequent analyses and modelling efforts.

3.4 Feature Extraction

In the ML workflow, feature extraction occurs after feature selection. A feature selection and classification process using a RandomForestClassifier on a dataset. The dataset is loaded and split into training and testing sets. Initially, a RandomForestClassifier is trained on the entire feature set, and feature importances are extracted. The top-k features are then selected based on their importances, and their names along with respective importances are printed. Following this, the SelectFromModel method is utilized to choose features based on a specified threshold. The transformed feature matrix is obtained, and a new RandomForestClassifier is trained on the selected features. The ultimate step involves evaluating the model's performance on the testing set, and key metrics such as the original number of features, the number of selected features, and the accuracy with the selected features are printed. This process helps identify the most influential features for classification while potentially reducing dimensionality. The accuracy score serves as a performance metric, reflecting how well the model performs with the selected features. This comprehensive workflow is valuable for enhancing model interpretability and efficiency in scenarios where feature selection is crucial.

In this feature selection and classification, the original dataset comprised 22 features after removing the column work_life_balance_score. Through a meticulous process involving a RandomForestClassifier, feature importances were assessed, and the top 10 features were selected based on their significance, which were matching the number of features selected in the highly correlated features. The accuracy evaluation of the model, utilizing this subset of features, yielded a commendable accuracy score of approximately 82.05%. This outcome signifies that the selected features, determined through their importance in the classification task, capture essential information for accurate predictions. The reduction from 22 to 10 features not only enhances computational efficiency but also demonstrates the effectiveness of feature selection in maintaining or even improving model performance. The accuracy metric provides valuable insight into the reliability and predictive power achieved with the streamlined set of features in the dataset.

3.5 Build Model

The next step in the ML process after feature extraction is to build a model. This research initially tested the build model with various ML algorithms such as classification and regression trees (CART), k-nearest neighbors (KNN), naive bayes (NB), random forest classifier (RFC), logistic regression (LR), LightGBM (LGB), support vector machine (SVM), extreme gradient boosting (XGB) and linear discriminant analysis (LDA) resulting in Table 4. The scores represent the performance of each algorithm on the dataset [28], with higher accuracy values indicating better performance at classification. The standard deviations provide a measure of the variability in the accuracy scores across different evaluations of the dataset. The preliminary comparisons made decisions to select algorithms that demonstrated higher accuracy, which are suitable choices for this research.

Table 4. Algorithms with mean accuracy and standard deviation

Algorithm	Mean Accuracy	Standard Deviation
CART (Classification and Regression Trees)	0.766	0.011
KNN (K-Nearest Neighbors)	0.854	0.012
NB (Naive Bayes)	0.888	0.009
RFC (Random Forest Classifier)	0.893	0.008
LR (Logistic Regression)	0.919	0.005
LGB (LightGBM)	0.936	0.005
SVM (Support Vector Machine)	0.943	0.005
XGB (Extreme Gradient Boosting)	0.942	0.005
LDA (Linear Discriminant Analysis)	0.975	0.004

The reported accuracy results represent the performance of various classification algorithms on dataset, evaluated through 10-fold StratifiedKFold cross-validation however only above 90% took into considerations. The LR demonstrates a solid performance, achieving a mean accuracy of approximately 91.87% with a low standard deviation of 0.49%. LGB follows closely, highlighting a competitive mean accuracy of around 93.57% and a slightly higher standard deviation of 0.53%. SVM exhibits robust performance, recording a mean accuracy of about 94.30% with a standard deviation of 0.50%. XGB closely trails SVM, securing a mean accuracy of approximately 94.22% and a slightly higher standard deviation of 0.55%. LDA outperforms the other algorithms, attaining a remarkable mean accuracy of 97.50% with a notably low standard deviation of 0.42%. These results provide valuable insights into the relative strengths and consistencies of each algorithm, aiding in the selection of an appropriate model based on 90% and above accuracy considerations for the dataset and classification task at hand.

3.6 Evaluate model

The next step in the ML process after building the model is to evaluate the model. The stacking classifier is a meta-ensemble approach that uses predictions from numerous base models to increase overall performance. The dataset is divided into two sets: training and testing, and the goal variable work_life_balance_score is created. The foundation models include the LGB classifier, SVM with radial basis function kernel, XGBoost classifier, and LDA. LR is the meta-model for aggregation. The stacking classifier is built using these basis models, and the final estimator is LR. The pipeline comprises preprocessing processes, especially Standard Scaling, which standardizes the features. The model is then trained using the training data, and predictions are produced on the test set. Following training, the model's performance is assessed using accuracy as a measure. The accuracy score is provided, offering a comprehensive assessment of the model's prediction performance. Furthermore, a full classification report is created, providing information on accuracy, recall, and F1-score for each class in the WLBS variable. These assessments help to provide a full knowledge of the stacking classifier's efficacy in predicting WLBS based on the provided characteristics.

Confusion Matrix: A confusion matrix is used in classification to determine the effectiveness of a ML method. It provides a summary of a classifier's predictions for a collection of data that has been verified. The matrix comprises four entries: true positive (TP), false positive (FP), true negative (TN), and false negative (FN). TP (true positives) is the number of instances correctly predicted as positive; TN (true negatives) is the number of instances correctly predicted as negative; FP (false positives) is the number of instances incorrectly predicted as positive; and FN (false negatives) is the number of instances wrongly predicted as negative. These data are used to construct a few indicators for assessing the model's performance. The accuracy is the proportion of correctly predicted instances to the total instances.

$$Accuracy = \frac{TP+TN}{TP+TN+\mathrm{FP}+\mathrm{FN}} \tag{1}$$

Precision is the ratio of correctly predicted positive observations to the total predicted positives.

$$Precision = \frac{TP}{TP+\mathrm{FP}} \tag{2}$$

Recall (also known as sensitivity or true positive rate) is the ratio of accurate predictions of positive observations to all observations in the class.

$$Recall = \frac{TP}{TP+\mathrm{FN}} \tag{3}$$

Recall is defined as the proportion of correctly predicted positive observations to total actual positives. Again, aid is used to provide context rather than performing the task itself. F1-Score is the weighted average of precision and recall.

$$F1 - Score = \frac{2\times Precision\times Recall}{Precision+Recall} \tag{4}$$

The F1 score represents the harmonic mean of accuracy and recall. Support, like accuracy and recall, is critical for understanding the whole context of model performance. Support is the number of examples or data points in each class in classification jobs. It is the number of observations belonging to a specific class. It is utilized in a variety of measures, including accuracy, recall, and F1-score, to contextualize a model's performance.

Table 5 summarizes the performance evaluation of an ensemble algorithm in predicting WLBS before undergoing feature scaling, selection, and extraction. The algorithm demonstrates an overall accuracy of 98.7%, highlighting its efficacy in classifying different WLB categories. The precision, recall, and F1-score metrics are presented for each WLBS class (0, 1, 2, 3), offering a detailed assessment of the algorithm's performance. Notably, the Ensemble Algorithm highlights high precision, recall, and F1-score across all classes, indicating its ability to accurately identify and classify instances within each WLB category. The macro average, which provides an overall performance metric across classes, emphasizes the algorithm's balanced performance. Similarly, the weighted average considers class imbalances, and in this case, it underscores the algorithm's exceptional performance, achieving high scores in precision, recall, and F1-score. The presented results suggest that the ensemble algorithm, in its pre-processed state, holds promise as a robust and accurate model for predicting WLBS.

Table 5. Performance evaluation on pre-feature processing

Algorithms	WLBS	Precision	Recall	F1-score	Support
Ensemble Algorithm **Accuracy: 0.987**	0	0.95	0.92	0.94	39
	1	0.99	0.99	0.99	1016
	2	0.99	0.99	0.99	1818
	3	0.96	0.95	0.95	185
	macro avg	0.97	0.96	0.97	3058
	weighted avg	0.99	0.99	0.99	3058

Table 6 provides a comprehensive overview of the performance evaluation for an ensemble algorithm after undergoing feature scaling, selection, and extraction. The algorithm achieves exceptional accuracy, with a perfect score of 100%, indicating its flawless ability to predict WLBS. Precision, recall, and F1-score metrics are presented for each WLBS class (0, 1, 2, 3), highlighting the algorithm's impeccable performance in accurately classifying instances within each WLB category. For each class, the ensemble algorithm attains a perfect precision, recall, and F1-score of 1.00, highlighting its capability to precisely identify and classify instances across all WLBS categories. The macro average, which offers an overall performance metric across classes, emphasizes the algorithm's consistent and outstanding performance. Similarly, the weighted average, considering class imbalances, reinforces the algorithm's exceptional accuracy and efficacy. The Ensemble Algorithm, post-feature processing, demonstrates unparalleled performance, achieving a perfect accuracy score and excelling in precision, recall, and F1-score metrics for each WLB category. These results suggest that the algorithm, with optimized features, is a robust and highly dependable model for predicting WLBS.

Table 6. Performance evaluation on post-feature processing

Algorithms	WLBS	Precision	Recall	F1-score	Support
Ensemble Algorithm **Accuracy: 1.00**	0	1.00	1.00	1.00	39
	1	1.00	1.00	1.00	1016
	2	1.00	1.00	1.00	1818
	3	1.00	1.00	1.00	185
	macro avg	1.00	1.00	1.00	3058
	weighted avg	1.00	1.00	1.00	3058

The performance evaluation on pre- and post-feature processing are shown in the following Figure 3.

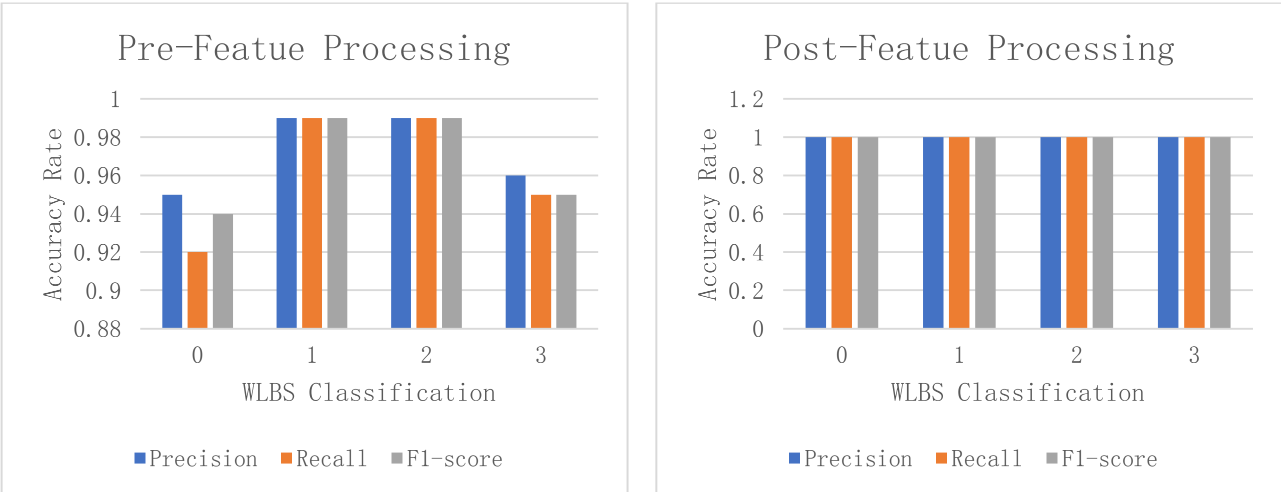

Fig. 3. Performance Evaluation on Pre- and Post-feature Processing

The predicted values of ensemble pre-feature and post-feature are shown in the following Figure 4, which are matching with the predicted scores. The scatter plot visually depicts the comparison between actual and predicted values generated by the Ensemble Algorithm post-feature processing. Each point on the plot represents an observation from the test dataset, with the x-axis indicating the actual values (ground truth) and the y-axis representing the corresponding predicted values. The points are color-coded according to the respective WLBS categories, providing a clear distinction between different classes.

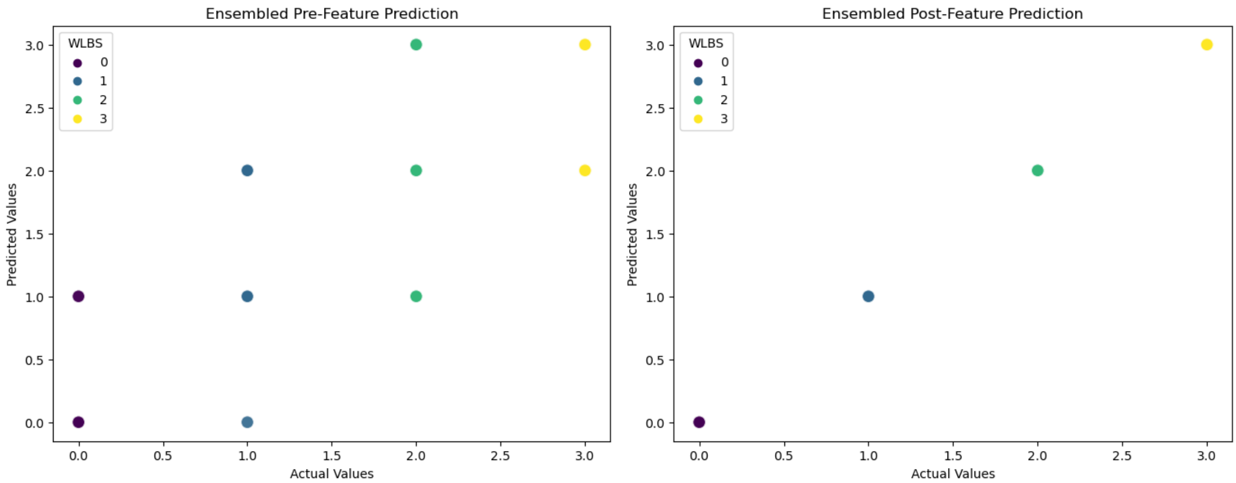

Fig. 4. Predicted Values of Ensemble Pre- and Post-Feature

The confusion matrix of ensemble pre-feature and post-feature is shown in the following Figure 5, which are matching with the predicted support scores. The generated confusion matrix and corresponding heatmap offer a comprehensive visual representation of the performance of the ensemble algorithm after feature processing. This matrix is a tabular summary of the algorithm's predictions compared to the actual labels from the test dataset. The x-axis represents the predicted labels, while the y-axis corresponds to the true labels. The heatmap is color-coded to highlight the distribution of correct and incorrect predictions, using varying shades of blue to indicate the intensity of the count in each cell. Annotations within the cells provide the numerical values of the confusion matrix, specifically indicating the count of observations falling into dissimilar categories.

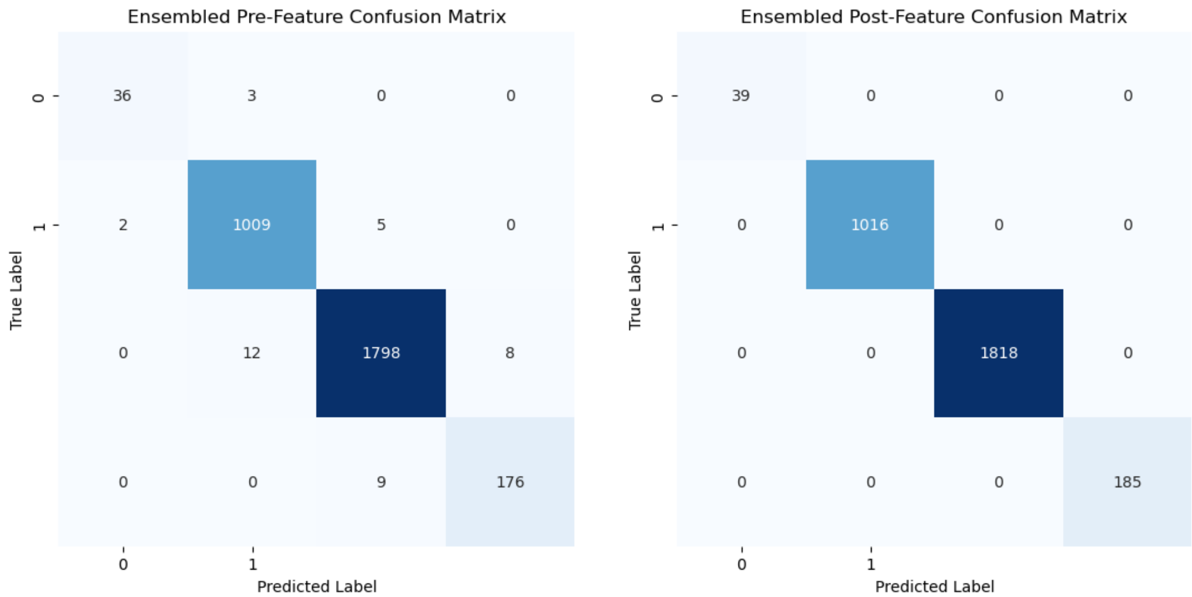

Fig. 5. Confusion Matrix of Ensemble Pre- and Post-Feature

3.7 *Cross-Validation*

The ultimate step in the ML process is to cross-validation to prevent overfitting, which occurs when a model is trained too well on the training data and performs poorly on new, unremarkable data. The provided performance evaluation results compare the model's accuracy before and after feature scaling, selection, and extraction using k-fold cross-validation and stratified k-fold cross-validation techniques. Before preprocessing, the k-fold cross-validation yields accuracy scores of approximately 90.79%, 91.94%, 91.94%, and 90.97%, respectively, for the WLBS. Meanwhile, the stratified K-fold cross-validation provides scores around 91.58%, 90.84%, 91.76%, and 91.08%, respectively, for the WLBS. After applying feature scaling, selection, and extraction, the model's performance significantly improves. Both k-fold cross-validation and stratified k-fold cross-validation consistently report perfect accuracy scores of 100% across all folds, represented as 1.0, 1.0, 1.0, and 1.0, respectively, for the WLBS. This remarkable enhancement suggests that the preprocessing steps have positively impacted the model's ability to accurately predict WLBS based on the given features. The perfect scores achieved after preprocessing indicate a prominent level of predictive accuracy and model robustness.

3.8 Comparison with other classifications

Table 7 presents a complete comparison of the performance assessment of two ML techniques, SVM and LR, in predicting WLBS, as illustrated in Figure 6. For the SVM model, the accuracy achieved is 97.93%. The precision, recall, and F1-score metrics are reported for each class (0, 1, 2, and 3), along with their respective support values. Notably, the model exhibits high precision, recall, and F1-score across different classes, demonstrating its effectiveness in distinguishing between WLB categories. Moving on to the LR model, it demonstrates an even higher accuracy of 99.67%. Like the SVM model, precision, recall, and F1-score metrics are reported for each class, as well as macro and weighted averages. The LR model highlights exceptional performance, achieving high scores in precision, recall, and F1-score across all classes. The macro and weighted average metrics further emphasize the overall robustness and accuracy of the model. In summary, both SVM and LR models exhibit outstanding predictive capabilities in determining WLBS, with the LR model slightly outperforming the SVM model in terms of accuracy. The detailed breakdown of precision, recall, and F1-score metrics provides a subtle understanding of the models performance across different WLB categories. While comparing these two models with ensemble classification, we found that the ensemble gave outstanding performance.

Table 7. Comparison of Performance Evaluation with other Classifications

Algorithms	WLBS	Precision	Recall	F1-score	Support
Support Vector Machines (SVM) **Accuracy: 97.93**	0	0.97	0.72	0.82	39
	1	0.97	0.98	0.98	1016
	2	0.98	0.99	0.99	1818
	3	0.98	0.91	0.94	185
	macro avg	0.97	0.90	0.93	3058
	weighted avg	0.98	0.98	0.98	3058
Logistic Regression (LR) **Accuracy: 99.67**	0	0.97	0.90	0.93	39
	1	0.99	1.00	1.00	1016
	2	1.00	1.00	1.00	1818
	3	1.00	0.99	0.99	185
	macro avg	0.99	0.97	0.98	3058
	weighted avg	1.00	1.00	1.00	3058

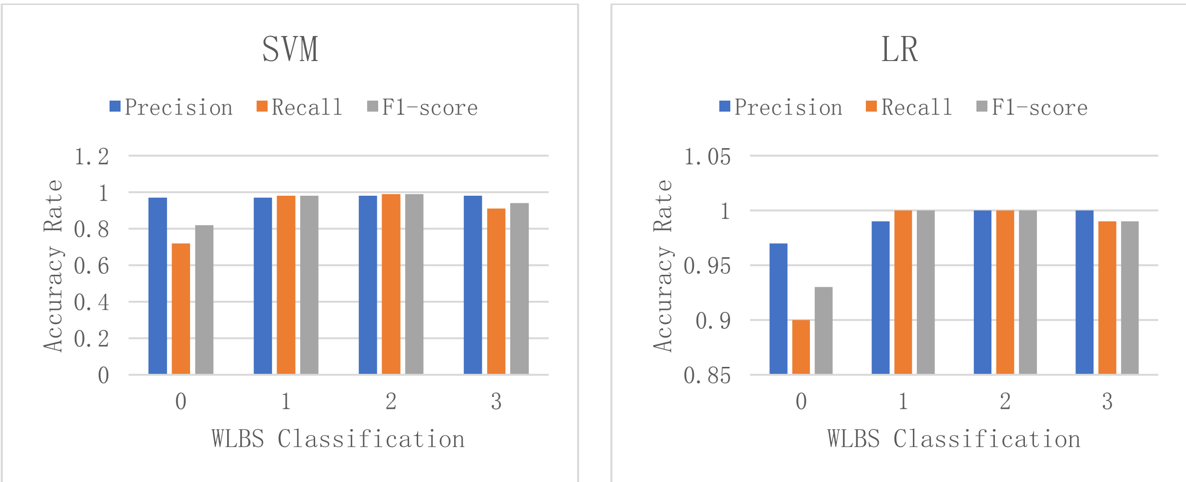

Fig. 6. Comparison of Performance Evaluation with other Classifications

The similar study [29] accuracy rate for Random Forest is 0.768733, for LR it is 1, and for Decision Tree it is 0.6551868. When these accuracy values are compared, LR has the highest accuracy with a value of 1, which may imply that the data was overfitted. Random Forest has a lesser accuracy of 0.768733, while Decision Tree has the lowest accuracy of 0.6551868. Another similar study [30], predict effective working schedules, increasing efficiency and promoting WLB. By analysing 12,756 people's data, 80% was trained using RFC, SVM, and NB algorithms. The algorithms predicted WLB with 71.5% accuracy, highlighting the importance of considering several factors in WLB. When compared to their model, the present research work showed excellent accuracy.

4. Exploratory Data Analysis (EDA)

This research used EDA to investigate and explore data sets to understand the key characteristics, find patterns, and establish correlations between variables, which are grouped into the lifestyle and WLB of professionals striving to develop a more balanced and meaningful life. The lifestyle is subdivided into healthy body and mind as well as happiness. Recalling Table 1, the following subplots show WLBS ranging from 0 to poor, 1 to average, 2 to good, and 3 to very good are the legends illustrated on the charts.

RQ1: Low Correlated Features (Life-Style): How do lifestyle factors, including daily steps, body mass index (BMI) range, daily sleep hours, weekly meditation, fruits and veggies consumption, daily shouting, daily stress, places visited, core circle, sufficient income, and lost vacation, contribute to the perceived work-life balance score (WLBS) across different age groups?

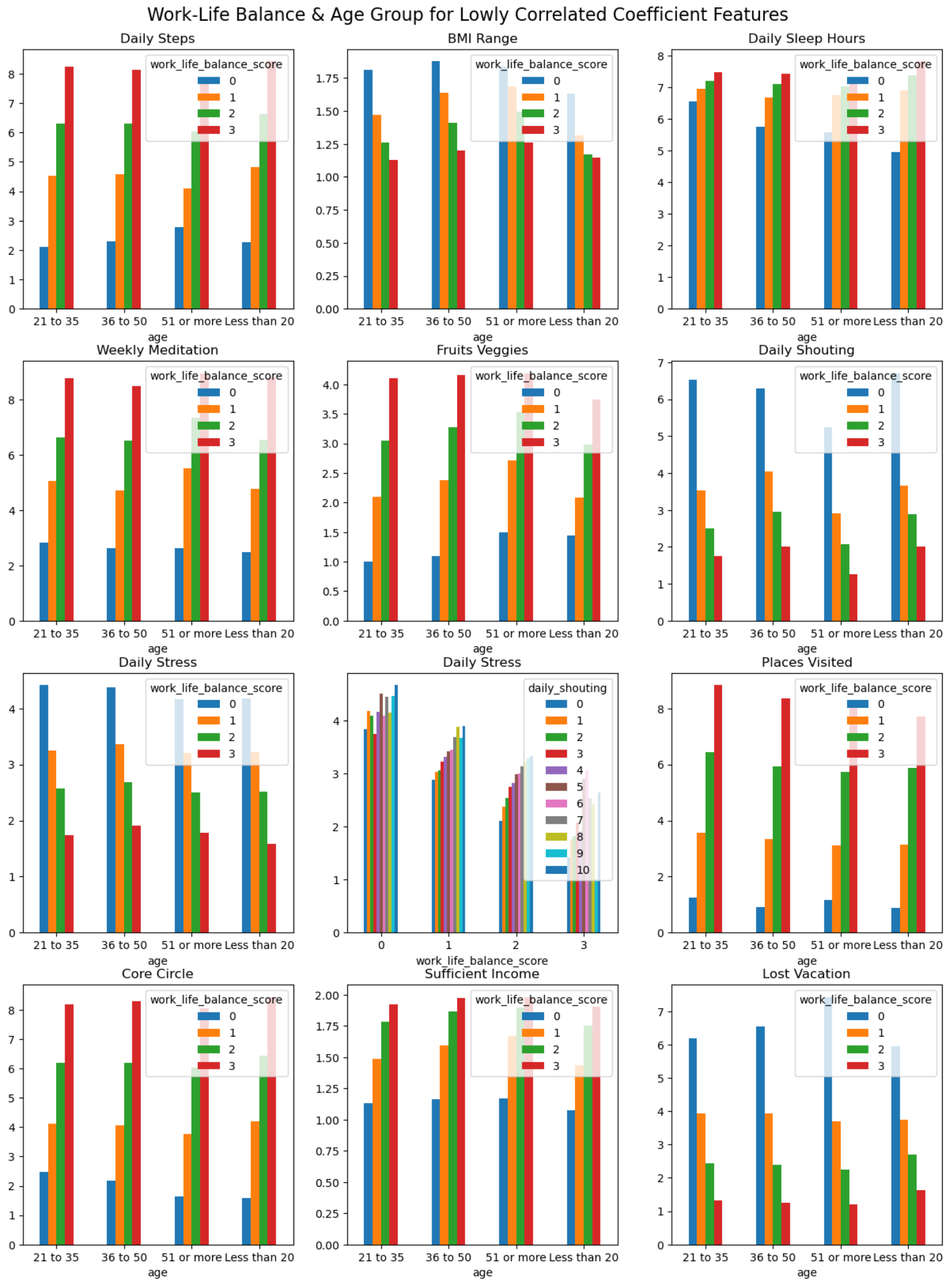

Fig. 7. Work-Life Balance & Age Group for Lowly Correlated Coefficient Features

Figure 7 created using Python code, provides a comprehensive set of visualizations within a 4x3 grid, each depicting the lowly correlated features between various health-related features and WLBS across different age groups. Each subplot corresponds to a distinct feature such as daily steps, BMI range, daily sleep hours, weekly meditation, fruits and vegetables, daily shouting, daily stress, place visited, core circle, sufficient income, and lost vacation, highlighting their relationship with WLB. Utilizing bar plots, the code leverages the Pandas library's pivot table functionality to organize and present the data effectively. The x-axis labels are thoughtfully rotated for improved readability. The overall layout is polished with a main title, 'Work-Life Balance and Age Group for Lowly Correlated Coefficient Features,' providing context to the entire visualization. By investigating these lowly correlated coefficient features, the research aims to understand the potential influence of these lifestyle variables on an individual's WLB, considering the subtle perspective brought about by age differences. The subplots present a detailed breakdown of each lifestyle factor's impact on WLB, providing a foundation for a comprehensive examination of how these factors interplay with age to shape individuals' perceptions of their WLB.

Increased daily routines, a lower BMI, and more sleep are found to improve WLBS. Similarly, increased weekly meditation and consumption of fruits and vegetables also contribute to WLBS. Daily shouting and stress levels are found to be detrimental to WLBS, with daily shouting being a significant factor. More places visited and a core circle provides moral support, enhancing WLBS. Sufficient income supports family needs and helps focus on WLB. Less vacation time is associated with poor WLBS, with less vacation time contributing to better WLB. The daily steps, sleep hours, weekly meditation, fruits and vegetables, places visited, core circle, and sufficient income features are supportive in a positive way to WLB, and the BMI range, daily stress, daily shouting, and lost vacations are negative to health and affect WLB. In conclusion, a balanced WLB is crucial for overall well-being. Increased daily steps, lower BMI, increased sleep, and a core circle provide moral support, while sufficient income supports family needs and helps focus on WLB. The vacation can also contribute to better WLBS and bring mindfulness. Overall, a well-balanced WLB is essential for overall well-being.

RQ2: High Correlated Features (Work-Life): What is the impact of specific lifestyle and work-related aspects, such as time for passion, social network engagement, donation activities, supporting others, experiencing a state of flow, completing daily tasks (to-do completed), achieving personal goals, receiving personal awards, and having a clear life vision, on the perceived work-life balance score (WLBS) within various age groups?

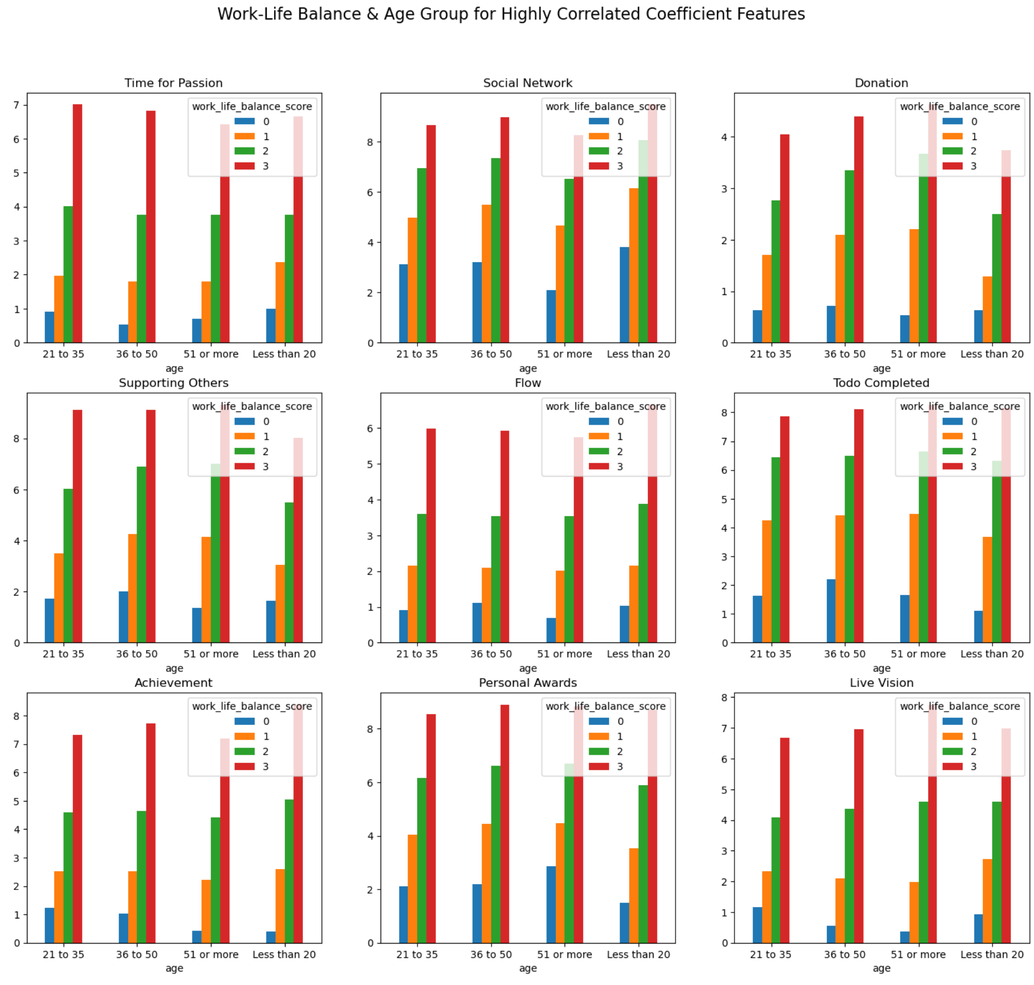

Fig. 8. Work-Life Balance & Age Group for Highly Correlated Coefficient Features

Figure 8 created using Python code, generates a visually compelling 3x3 grid of bar plots, illustrating the impact of various personal and motivational factors on WLB across different age groups. Each subplot focuses on a distinct feature such as time for passion, social network, donation, supporting others, flow, to-do completed, achievement, personal awards, and live vision. Leveraging the Pandas library's pivot table functionality, the code organizes and presents the data effectively, utilizing bar plots to highlight the highly significant correlation with WLBS. The x-axis labels are carefully adjusted for readability. The overall layout is aesthetically pleasing, with a main title, 'Work-Life Balance and Age Group for Highly Correlated Coefficient Features,' providing context to the entire visualization. By scrutinizing these highly correlated features, the research aims to uncover the distinct patterns and associations between these factors and WLB, considering age-specific perspectives. The subplots provide a comprehensive visual representation of how each of these features influences WLB across different age cohorts, setting the stage for a detailed investigation into the dynamics of these elements in shaping individuals' perceptions of their WLB.

The highly influenced features are social networks, personal awards, time for passion, supporting others, achievement, donations, to-dos, flow, and a live vision. All the features are shown as very good WLBS, followed by good, then average, and least poor. This means that all the features are supported positively and there are no negative features. When we observe the above 51 age groups in all donations, supporting others, and completing tasks, personal awards and live vision is high and poor in time for passion, social networks, flow of work, and achievement. When we observe the below 20 age groups in all donations, supporting others is poor, and time of passion, social networks, the flow of work, to-do completed, achievement, personal awards, and live vision are high. When we observe the age group of 21 to 50, they are moderately high in all the features.

RQ3: Work-life balance (WLB): What is the prevalent distribution of work-life balance score (WLBS) among the surveyed individuals, and how does this distribution reflect on the overall perception of work-life balance (WLB) within the studied population?

Figure 9 is an informative pie chart representing the distribution of WLBS across dissimilar categories. The four labels, 'Poor,' 'Average,' 'Good,' and 'Very Good,' categorize the data into distinct segments. The sizes of each segment are proportional to the corresponding values in the WLB group list, which represents the count or frequency of individuals falling into each WLB category. The percentage contribution of each category is displayed on the chart for quick interpretation, which is 58.1% very good, 34.5% average, 6.1 good, and 1.3 poor. The use of distinct colours, such as red, orange, blue, and green, enhances visual distinction between the categories. The chart's central title, 'Work-Life Balance Score,' provides clear context, making it easy for viewers to grasp the overall distribution of WLBS in a visually compelling manner.

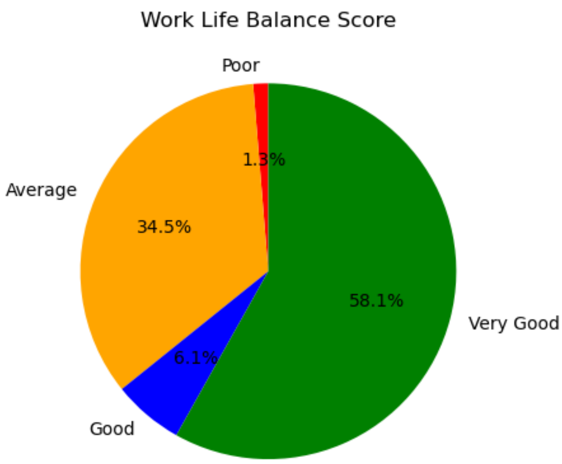

Fig. 9. Percentage of WLBS in Categories

RQ4. Correlation factors: What are low (Life-Style) and high (Work-Life) and selected correlated coefficient features?

The threshold value is set at.30, which is filtered with low and high values for the correlation coefficient. The low correlation matrix provided offers insights into the relationships between various features, presenting correlation coefficients that quantify the strength and direction of associations. Each cell in the matrix represents the correlation between two features. Notably, the correlation between fruits and vegetables and daily stress is low and negative (-0.097334), implying that an increase in the consumption of fruits and vegetables correlates with a slight decrease in daily stress levels. Similarly, fruits, veggies, and places visited exhibit a low positive correlation (0.255699), indicating that those who consume more fruits and vegetables tend to visit more places. This pattern persists across several pairs of features, such as the core circle, BMI range, daily steps, sleep hours, lost vacation, daily shouting, sufficient income, weekly meditation, age, and gender. In each case, the correlation is low and either positive or negative, signifying weak linear relationships between the variables. The understanding of these correlation coefficients provides valuable insights into the potential interplay between lifestyle factors and habits, as shown in Figure 10a.

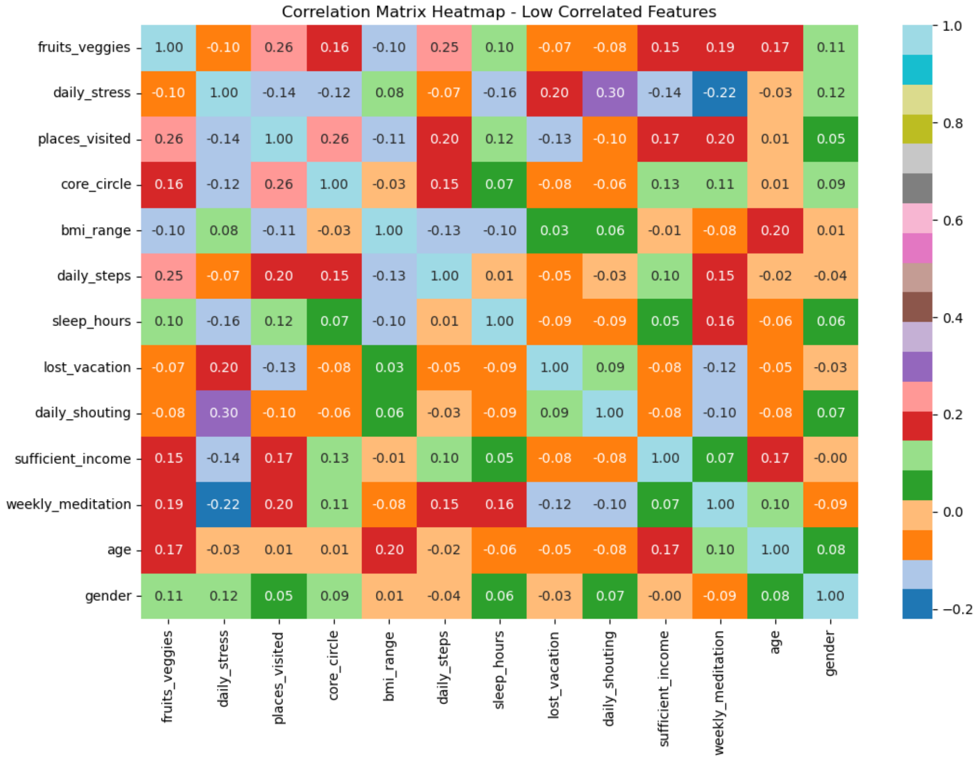

Fig. 10a. Lowly Correlated Features

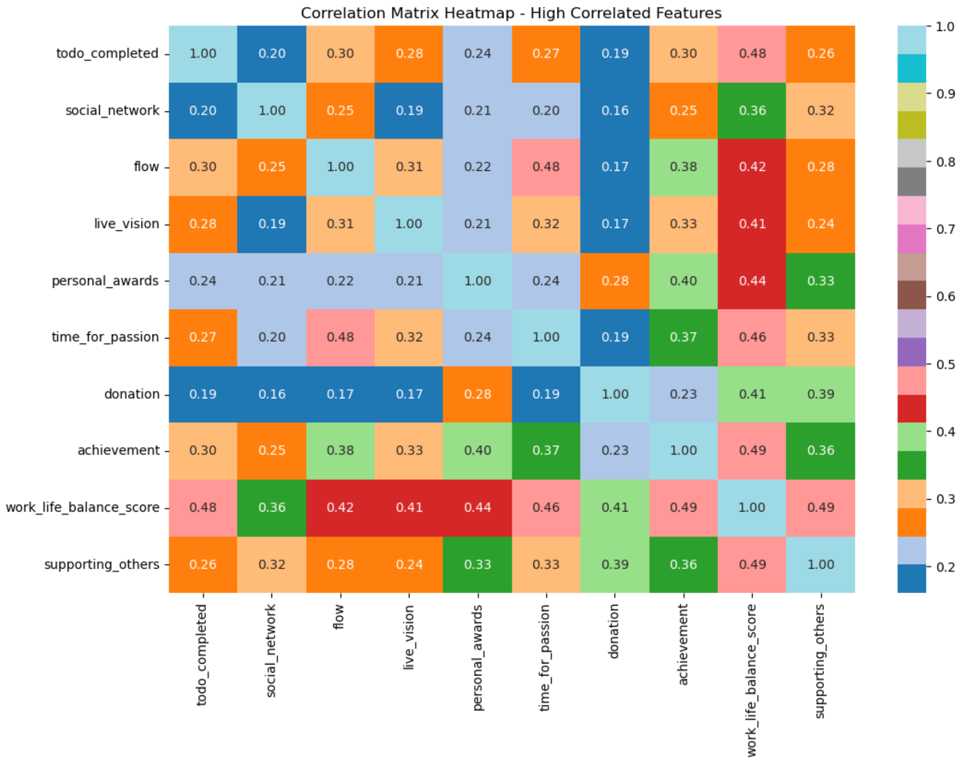

Fig. 10b. Highly Correlated Features

The correlation coefficients presented in the above figure reveal significant associations between various features, and the identification of highly correlated features is paramount for understanding the potential interdependencies shown in Figure 10b. Notably, personal awards and achievement exhibit a correlation coefficient of 0.396326, signifying a moderately positive correlation. This suggests that individuals honoured with personal awards are more likely to achieve higher levels of success. Similarly, personal awards and WLBS show a correlation coefficient of 0.436202, indicating a moderately positive correlation. This implies that those who have received personal awards tend to experience a better WLB. Another notable correlation emerges between achievement and time for passion, with a coefficient of 0.367496, suggesting that individuals with elevated achievement levels are more inclined to dedicate more time to pursuing their passions. Lastly, the WLBS and supporting factors reveal a correlation coefficient of 0.485233, indicating a moderately positive correlation. This shows that those with better WLBS are more inclined to participate in activities that benefit others. It is critical to highlight that correlation does not indicate causation, and more in-depth investigation is required to determine any causal linkages between these traits.

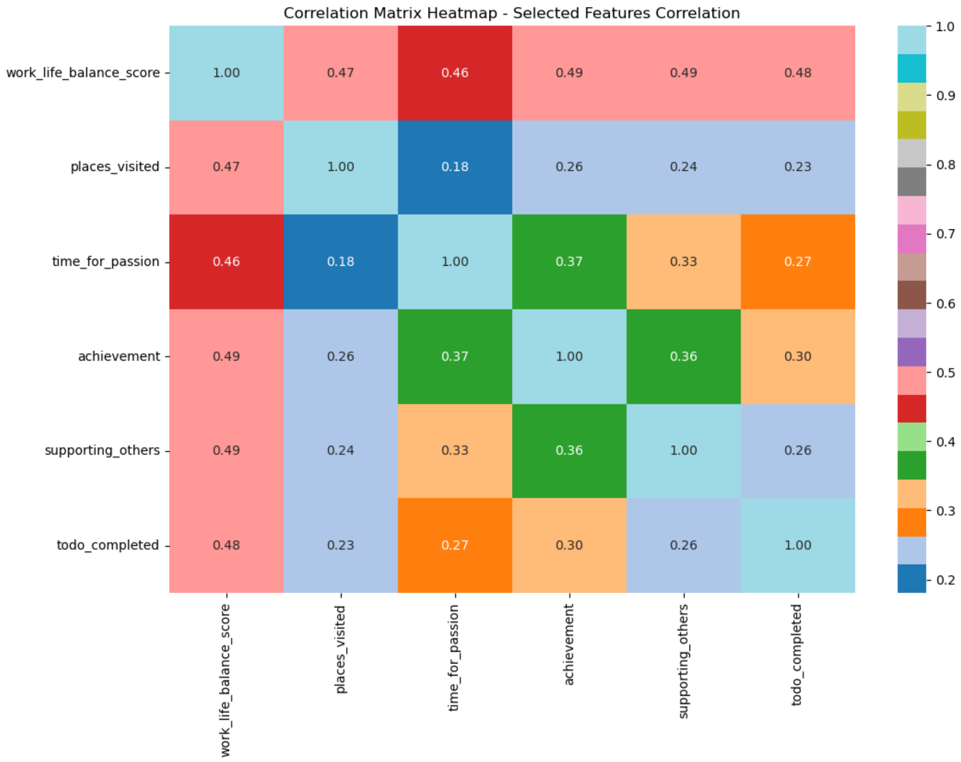

Fig. 11. Selected Features Correlation

The selected features of correlation provide insights into the relationships between different variables, unveiling the strength and direction of the associations shown in Figure 11. Specifically, the analysis focuses on the WLBS and its correlations with various aspects of an individual's life. The positive correlation of 0.468439 between WLBS and places visited indicates that individuals with a higher WLBS tend to explore and visit more places. Similarly, the positive correlation of 0.457858 with time for passion suggests that those with a better WLB allocate more time to their passions. Notably, the correlation of 0.492228 between WLBS and achievement implies that individuals achieving their goals are more likely to have a higher WLBS. Additionally, positive correlations of 0.485233 and 0.483263 with supporting others and to-do completed, respectively, signify that a superior WLB is associated with increased engagement in supporting others and accomplishing tasks. It is crucial to note that the correlation coefficients range from -1 to 1, with 1 representing a strong positive correlation, -1 representing a strong negative correlation, and 0 indicating no correlation. These findings underscore the interconnectedness between WLB and various aspects of personal and professional life.

RQ5: Work-life balance (WLB) and holistic well-being (HWB): What is the interplay between work-life balance (WLB) and holistic well-being (HWB)?

A robust state of well-being encompasses both physical and mental health, where daily steps and weekly meditation contribute to bodily wellness, while exploring unfamiliar places and dedicating time to passions nurture mental well-being. The assessment of WLB is multifaceted, considering completion of tasks and achievements, which in turn influence personal recognition and alignment with one's life vision as depicted in Figure 12. Each subplot within the generated figure illustrates a distinct facet of well-being in relation to corresponding WLBS. The first subplot describes the correlation between daily steps and weekly meditation against WLBS, while the second subplot elucidates the interplay between places visited and time allocated for passions. In an equivalent manner, the third subplot interprets the connection between achievement levels and task completion concerning WLB, whereas the fourth subplot explores personal awards in association with one's life vision. Overall, all subplots have a WLBS of 3, which is very good. Through these visualizations, it facilitates a comprehensive understanding of the complicated dynamics between WLB and different proportions of well-being.

The research investigated the correlation between lifestyle and WLB, using machine learning techniques to predict WLB through a dataset. The results demonstrated that the ensemble algorithm achieved a perfect accuracy score of 100% in predicting WLBS, while the LR model had an accuracy of 99.67%. The research demonstrates the effectiveness of ML in understanding and predicting WLB. However, it is significant to remark that the research had limitations, such as using a pre-existing dataset and not gathering real-time data. The findings may not be accurate to all industries and organizational contexts. This research provides valuable insights into the complex dynamics of WLB and provides practical recommendations for individuals and organizations striving to achieve a more balanced and meaningful life. Research suggests that ML can enhance WLB and overall health by promoting flexibility, personalization, and continuous evaluation.

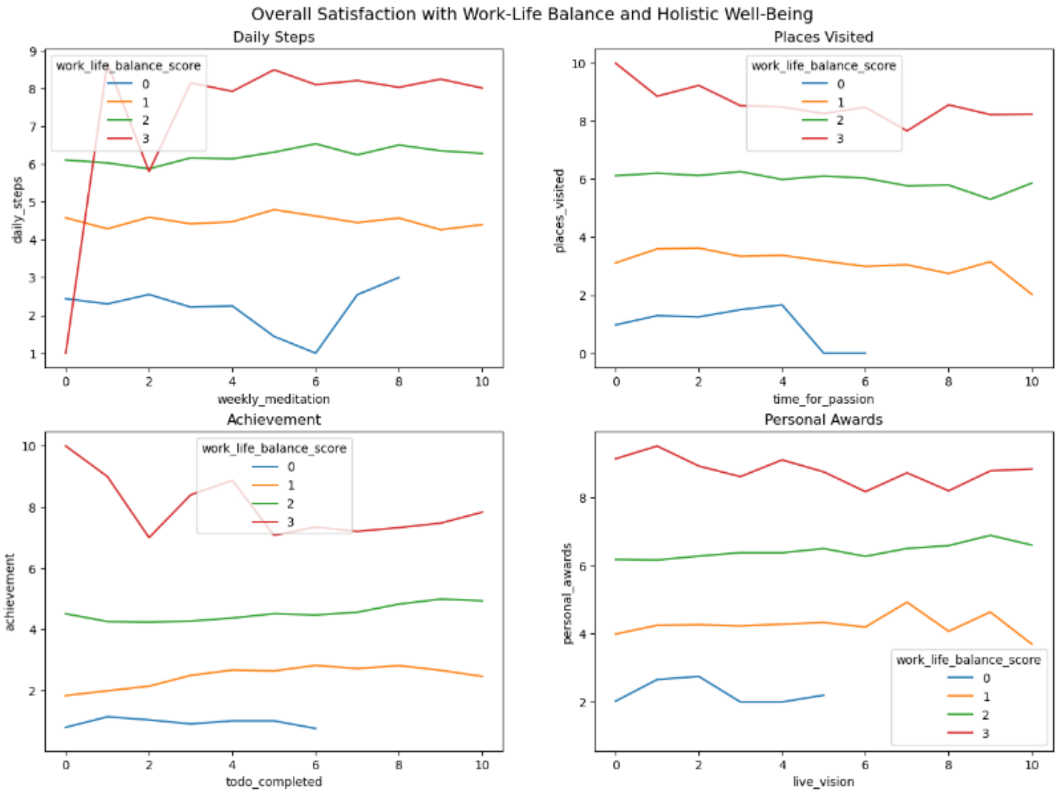

Fig. 12. Selected Features Correlation

The following recommendations can be tailored to specific organizational or policy contexts, highlighting the need for ongoing research and evaluation to optimize the use of ML in this area. However, these recommendations should be tailored to individuals' needs and specific organizational circumstances.

- Conduct regular assessments to evaluate the effectiveness of interventions and programs aimed at improving WLB.
- Design and implement employee wellness programs to address physical health, mental health, and wellbeing.
- Encourage individuals to allocate time for their passions and personal goals, aiding with attain goals.
- Foster a culture of WLB by initiating an encouraging work environment and promoting WLB policies.
- Foster a supportive workplace to encourage collaboration, teamwork, and mutual support among employees.
- Implement flexible work arrangements options to flexible working hours and enhance WLB and wellbeing.
- Implement wellness initiatives based on ML insights to address specific areas of concern.
- Promote work-life harmony by promoting employees to balance work obligations and personal desires.

5. Conclusion

This research investigated the complex interplay between lifestyle choices and work-life balance, aiming to provide a holistic perspective on individual well-being. By extending conventional well-being concepts, the research leverages machine learning algorithms and correlation analyses to unravel the relationships between various features. The correlation coefficients highlight significant associations, providing valuable insights into how different aspects of life influence work-life balance. Supervised learning algorithms, including ensemble methods, are employed to predict work-life balance scores, showcasing promising accuracy and robustness. Exploratory data analysis enhances our understanding of the dynamic interactions between lifestyle factors and work-life balance. The comparative analysis of models and preprocessing techniques emphasizes the importance of feature scaling, selection, and extraction in improving predictive accuracy. Notably, the ensemble algorithm emerges as a powerful tool for predicting work-life balance scores. The performance evaluation, both pre-feature and post-feature processing, demonstrates the efficacy of the ensemble algorithm, achieving perfect accuracy. The scatter plot and confusion matrix visually reinforce the algorithm's capability to accurately predict work-life balance scores. Further insights are provided by comparing the performance of support vector machines and logistic regression models, with both exhibiting high accuracy but the ensemble algorithm standing out. The study finishes in a comprehensive examination of the factors influencing work-life balance, revealing patterns across different age groups and shedding light on positively and negatively correlated features. The pie chart visually presents the distribution of work-life balance scores, emphasizing the prevalence of very good and average categories. Finally, the correlation matrix and selected features highlight the subtle relationships between variables, offering valuable implications for fostering a balanced and meaningful life. This research contributes to the ongoing debate on holistic wellbeing and provides actionable insights for individuals and professionals striving to achieve a more balanced and fulfilling life.

References

[1] Sharkey, J. (2019). Work-Life Balance versus Work-Life Merge: Rationale for Research. *DBS Business Review*, 03, 49–72.

[2] Sutton, A. (2020). Living the good life: A meta-analysis of authenticity, well-being and engagement. *Personality and Individual Differences*, 153, 109645. https://doi.org/10.1016/j.paid.2019.109645

[3] Gragnano, A., Simbula, S., & Miglioretti, M. (2020). Work–life balance: weighing the importance of work–family and work–health balance. *International Journal of Environmental Research and Public Health*, 17(3), 9–11. https://doi.org/10.3390/ijerph17030907

[4] Hasan, Z. U., Khan, M. I., Butt, T. H., Abid, G., & Rehman, S. (2020). The balance between work and life for subjective well-being: A moderated mediation model. *Journal of Open Innovation: Technology, Market, and Complexity*, 6(4), 1–18. https://doi.org/10.3390/joitmc6040127

[5] Dhandayuthapani, B. V. (2024). Exploring the Intersections of Responsible AI for Community Wellbeing in Local Governance: A Comprehensive Review, *i-manager's Journal on Humanities & Social Sciences, 4(1), 57-64. https://doi.org/10.26634/jhss.4.1.20327*.

[6] Mensah, A., & Adjei, N. K. (2020). Work-life balance and self-reported health among working adults in Europe: A gender and welfare state regime comparative analysis. *BMC Public Health*, 20(1), 1–14. https://doi.org/10.1186/s12889-020-09139-w

[7] Badri, M., Al Khaili, M., Aldhaheri, H., Yang, G., Al Bahar, M., & Al Rashdi, A. (2022). Examining the structural effect of working time on well-being: Evidence from Abu Dhabi. *Social Sciences and Humanities Open*, 6(1), 100317. https://doi.org/10.1016/j.ssaho.2022.100317

[8] Borowiec, A. A., & Drygas, W. (2023). Work–Life Balance and Mental and Physical Health among Warsaw Specialists, Managers, and Entrepreneurs. *International Journal of Environmental Research and Public Health*, 20(1). https://doi.org/10.3390/ijerph20010492

[9] Althammer, S. E., Reis, D., van der Beek, S., Beck, L., & Michel, A. (2021). A mindfulness intervention promoting work–life balance: How segmentation preference affects changes in detachment, well-being, and work–life balance. *Journal of Occupational and Organizational Psychology*, 94(2), 282–308. https://doi.org/10.1111/joop.12346

[10] Palumbo, R. (2022). Involved at work and disinvolved out of work: unraveling the implications of involvement on accountants' work–life balance. *Management Decision*, 61(13), 26–53. https://doi.org/10.1108/MD-03-2022-0345

[11] South Africa. (2022). Employee wellness programmes for improved performance, 1–12.

[12] Narayan, R. (2020). Leveraging Digital Intelligence for Community Well-Being, *International Journal of Community Well-Being*,539–558. https://doi.org/10.1007/s42413-020-00085-4

[13] Putra, A. S. B., Kusumawati, E. D., & Kartikasari, D. (2023). Unpacking the Roots and Impact of Workplace Well-being: A Literature Review. *International Journal of Multidisciplinary Approach and Research in Science*, 2(01), 312–321. https://doi.org/10.59653/ijmars.v2i01.433

[14] Stanley, S., & Sebastine, A. J. (2023). Work-life balance, social support, and burnout: A quantitative study of social workers. *Journal of Social Work*, 23(6), 1135–1155. https://doi.org/10.1177/14680173231197930

[15] Kalalo, R. R. (2024). The Influence of Work Life Balance on Employee Engagement with Work Motivation as A Moderating Variable, *Innovative: Journal of Social Science*, Vol 4, No. 1, 1506–1514.

[16] R. Phillips et al., "Introduction to the Special Issue: Intersections of Artificial Intelligence and Community Well-Being," *Int. J. Community Well-Being*, vol. 3, no. 4, pp. 425–435, 2020, doi: 10.1007/s42413-020-00099-y.

[17] D. Fernandez and S. Lebin, "The Relationship Between Work-Life Balance, Job Satisfaction, Age, and Gender," *J. Adv. Educ. Pract.*, vol. 3, no. 1, pp. 1–12, 2022, [Online]. Available: . https://openriver.winona.edu/jaep/vol3/iss1

[18] P. Susanto, M. E. Hoque, T. Jannat, B. Emely, M. A. Zona, and M. A. Islam, "Work-Life Balance, Job Satisfaction, and Job Performance of SMEs Employees: The Moderating Role of Family-Supportive Supervisor Behaviors," *Front. Psychol.*, vol. 13, no. June, pp. 1–12, 2022, doi: 10.3389/fpsyg.2022.906876.

[19] P. Thriveni, L. Benedicta, and D. Souza, "Work-Life Balance and Job Satisfaction of Higher Education Teachers-A Systematic Literature Review," *EPRA Int. J. Res. Dev.*, vol. 7838, no. January, pp. 89–100, 2023. https://doi.org/10.36713/epra2016

[20] S. Kedia, A. Kumar Shukla, and S. Sinha, "Balancing Act: A Comprehensive Study of Work-Life Balance and Its Impact on Employee Performance and Job Satisfaction," *Eur. Chem. Bull*, vol. 2023, no. 12, pp. 668–678, 2023, doi: 10.48047/ecb/2023.12.si12.057.

[21] A. Dahar, "Analyzing Professional Well-being: The Intersection of Demographics, Work-Life Balance and Job Satisfaction," Pakistan J. Med. Dent., no. January, pp. 88–95, 2024, doi: 10.36283/pjmd13-1/016.

[22] R. Tri, H. Safariningsih, J. P. Soehaditama, and D. Ulfah, "Employee Wellbeing : Analysis Work Life Balance , Resilience , Psychological Structure and Organizational Culture," East Asian J. Multidiscip. Res., vol. 3, no. 1, pp. 309–318, 2024.

[23] E. Gülmez, H. I. Koruca, M. E. Aydin, and K. B. Urganci, "Heuristic and swarm intelligence algorithms for work-life balance problem," Comput. Ind. Eng., vol. 187, no. December 2023, p. 109857, 2024, doi: 10.1016/j.cie.2023.109857.

[24] E. Kurtuluş, H. Yıldırım Kurtuluş, S. Birel, and H. Batmaz, "The Effect of Social Support on Work-Life Balance: The Role of Psychological Well-Being," Int. J. Contemp. Educ. Res., vol. 10, no. 1, pp. 239–249, 2023, doi: 10.33200/ijcer.1152323.

[25] A. Y. A. Bakar, “The Application of Work-Life Balance in Relation with Employees Well-Being: An Exploratory Study,” J. Law Sustain. Dev., vol. 12, no. 2, p. e3243, 2024, doi: 10.55908/sdgs.v12i2.3243.
[26] L. Qi, C. M. Yee, B. Chan, and Y. Fah, “The Role of Work-Life Balance in Enhancing Employee Loyalty,” *Account. Corp. Manag.*, vol. 6, pp. 43–47, 2024, doi: 10.23977/acccm.2024.060106.
[27] M. Izharuddin, “Job Satisfaction , and Working Environment : Mediating Role of Work-Life Balance Kepuasan Kerja dan Lingkungan Kerja : Peran Mediasi Work-Life Balance,” *Manag. Stud. Entrep. J.*, vol. 5, no. 1, pp. 1144–1149, 2024.
[28] Lifestyle & Wellbeing Dataset, Kaggle. Ret. Dec, 2023, from https://www.kaggle.com/code/carlyhall/lifestyle-and-wellbeing
[29] Alshanqiti, A., et al. (2021). WQD 7004 - Predicting the Work Life Balance Score. Retrieved from https://rpubs.com/IzzahR/1055496
[30] Radha, K., & Rohith, M. (2021). An Experimental Analysis of Work-Life Balance Among the Employees Using Machine Learning Classifiers. *International Journal of Computer Trends and Technology*, 69(4), 39–48. https://doi.org/10.14445/22312803/IJCTT-V69I4P108

Authors’ Profiles

Bala Dhandayuthapani V. is presently working as an IT faculty member at the University of Technology and Applied Sciences, Shinas campus, North Al Batinah, Sultanate of Oman. He has 22 years of experience as a faculty member, including in India, Ethiopia, and Oman. He received his PhD in information technology and computer science from Manonmaniam Sundaranar University, India. He received an M.Tech. in information technology from Allahabad Agricultural Institute of Deemed University, an M.S. in information technology, and a B.Sc. in computer science from Bharathidasan University. He has published forty peer-reviewed technical research papers in various international journals and conference proceedings. He has also written a textbook entitled "An Introduction to Parallel and Distributed Computing through Java." He has given different invited technical talks and has been involved in a wide range of academic activities.

How to cite this paper: Bala Dhandayuthapani V., "Exploring Holistic Well-Being: EDA and a Comparative ML Approach to Work-Life Balance", International Journal of Modern Education and Computer Science(IJMECS), Vol.16, No.6, pp. 1-19, 2024. DOI:10.5815/ijmecs.2024.06.01