Utilization of Machine Learning Techniques in Predicting Burnout in Organizations

Ifeoluwa Oduwaiye
Department of Computer Engineering, Bells University of Technology, Lagos Nigeria

Samuel Oyefusi
Worcester Polytechnic Institute, Department of Robotics Engineering,
Worcester, Massachusetts

Ayoola Okunlola
Department of Computer Engineering, Bells University of Technology, Lagos Nigeria

Melody Alabi
Department of Computer Engineering,
Bells University of Technology, Lagos Nigeria

ABSTRACT
Employee burnout due to high stress levels is one of the factors that impacts the productivity of employees as well as the quality of their output at work. It is therefore extremely important for HR personnel and management staff to predict and manage stress levels ahead of burnout. Traditional methods of identifying burnout often rely on retrospective assessments, limiting organizations' ability to address it proactively. Machine Learning has been effective in performing predictive analytics on data and has been applied by researchers in similar use cases such as employee attrition prediction, applicant scoring using AI-enhanced ATS systems, and employee performance analysis. This research aims to apply descriptive analysis techniques in analyzing the causes of burnout rates in organizations and will also utilize select machine learning algorithms such as Logistic Regression, Decision Trees, Random Forest, Support Vector Classifier, Gradient Boosting, Extreme Gradient Boosting, Stacking Classifier and Neural Networks in building a predictive model that can predict employee burnouts in organizations and flag these predicted high stress levels with specified personnels for monitoring. This research uses primary data gotten from a wide variety of professionals, which is used to train a machine learning model that is deployed to a web endpoint where it can be accessed by everyone. This work explores the potential of data-driven approaches in supporting employee well-being and offers actionable insights for organizations aiming to mitigate burnout. Future work may focus on expanding data sources and refining predictive models to enhance scalability and applicability in diverse organizational contexts.

Keywords
Employee Burnout, Machine Learning, Organizational Psychology, Mental Health

1. INTRODUCTION
1.1 Background of Study
In the modern workplace, characterized by rapid technological advancements, increased competition, and evolving demands, employee well-being has become a focal point for organizations seeking to maintain productivity and foster long-term success. (Bakker & Demerouti, 2018). Among the most pressing issues affecting the workforce today is employee burnout, a phenomenon defined by chronic physical, emotional, and mental exhaustion due to prolonged exposure to stressors in the workplace (Z. Liu et al., 2023). Burnout affects an individual's ability to perform their job effectively, leading to serious health consequences, a decline in organizational morale, and decreased job satisfaction (Spector, 2022), (Wahab & Anwar, 2023). Employee burnout poses significant challenges to organizations in various forms, such as reduced productivity, higher turnover rates, absenteeism, and diminished team collaboration. Burnout was officially recognized as an occupational occurrence by (McCarthy, 2025). This research further emphasized its global prevalence and impact in organizations today. Burnout causes a decline in motivation, loss of talent, recruitment expenses, and reduced overall efficiency at the workplace. Companies have been increasingly motivated to find solutions that pre-emptively address this issue.

Traditional approaches to managing burnout often rely on qualitative assessments, such as periodic employee surveys and observations by managers. (Huebner & Zacher, 2021). While these methods can provide some insights, they are limited by their reactive nature, subjectivity, and infrequent application. Such limitations make it difficult for organizations to detect early signs of burnout, especially in large workforces where individual attention may be constrained. Furthermore, the subjective interpretation of survey responses can lead to biased outcomes, further complicating efforts to implement effective burnout mitigation strategies. The integration of Machine Learning (ML) into human resource management presents a powerful, data-driven solution to the problem of employee burnout. ML algorithms can analyze vast amounts of structured and unstructured data, uncovering complex patterns that human analysis might overlook. The algorithm leverages data such as work hours, task completion rates, employee engagement scores, communication patterns, and even behavioral indicators from company-provided digital tools. ML models can predict burnout risk with high accuracy and provide insights that inform targeted intervention (Răducu & Stănculescu, 2022). Unlike traditional burnout assessments, ML enables continuous monitoring and real-time analysis, which significantly enhances an organization's ability to take proactive measures and prevent burnout before it fully manifests (Tehrani et al., 2022).

Several ML techniques, including supervised learning models like decision trees, support vector machines, and gradient boosting algorithms, have demonstrated potential in predicting burnout by processing relevant data features. (Ghavidel & Pazos, 2025). Unsupervised learning approaches, such as clustering and anomaly detection, can also identify typical patterns in employee behaviour that signal stress or disengagement. (Y. Liu et al., 2023). The choice of features and models plays a crucial role in the accuracy and reliability of burnout prediction systems. Identifying the most relevant data, whether drawn from time management tools, employee surveys, project management software, or direct performance metrics, ensures that predictive models yield actionable results. Moreover, the ethical considerations of deploying ML for burnout prediction cannot be overlooked. Data privacy, informed consent, and transparency in how data is utilized are essential to maintaining employee trust and safeguarding against misuse. Organizations must strike a balance between using advanced analytics to support employees and respecting their privacy and autonomy.

This paper examines the current state of ML for predicting employee burnout, exploring the types of data that contribute to reliable models, the benefits and limitations of various algorithms, and best practices for implementation. By integrating ML into burnout management strategies, organizations can foster a work environment that prioritizes employee well-being, ultimately contributing to higher retention, improved performance, and sustainable growth.

1.2 Research Aims and Objectives

This research aims to examine the effectiveness and applicability of machine learning algorithms in predicting employee burnout within organizational settings. The research objectives are:

1. To perform a comprehensive exploratory data analysis, identifying patterns, trends, and key factors contributing to employee burnout.
2. To develop and train multiple machine learning algorithms on employee burnout data, and select optimal models based on defined performance metrics.
3. To design and implement an accessible web application using Streamlit, enabling employees and employers to assess burnout risk by engaging with the predictive model.

2. LITERATURE REVIEW

2.1 Theories of Burnout

Burnout among average working employees has become an issue of pressing urgency, necessitating immediate attention and action. As work demands increase and the boundaries between personal and professional life blur, employees are more susceptible to chronic stress, which, if left unaddressed, escalates into burnout. Burnout is a prolonged stress effect on the body, which can lead to mental breakdown, delay in production, high turnover, and absenteeism. The term 'burnout' is similarly used with 'stress,' 'work strain,' and 'mental exhaustion.' As studies imply, burnout often can be the aftermath of stress, overload in work, lack of support, and resources. Burnout has also been strongly associated with reduced self-regulation capacity and overall well-being. (Gagnon et al., 2016). Research studies have indicated that job sectors like healthcare workers, emergency services, teachers, tech sectors, legal services, and civil services workers have a high rate of burnout. Employee burnout can be caused by different stressors, which will be discussed further. In this age of productivity, some professions would have a tendency towards burnout in employees, but creating limitations and management would reduce burnout and even employee turnover in organizations. Research studies have put into consideration how to point out factors, personalities, and health status that can be leading effects of employees, employers feeling overstressed, and detachment from jobs.

An early research done by (Maslach et al., 2001) . The impact of employee emotions and workplace relationships on burnout was studied, concluding that burnout is a state of overwhelming exhaustion, cynicism, and detachment in the job, feeling ineffective and unworthy of accomplishment. Maslach developed the three-dimensional model of burnout, which was formalized in the Maslach Burnout Inventory (MBI) (Maslach, 2013). This instrument gained popularity as the most common scale to gauge burnout (Enzmann et al., 1995; Shirom & Melamed, 2006). Burnout models are theories that are utilized to analyze, forecast, and prevent burnout, prolonged stress at the workplace that is not effectively dealt with. The Maslach Burnout Inventory (MBI), Job Demands-Resources (JD-R) Model, and Conservation of Resources (COR) Theory are some popular models of burnout.

Maslach Burnout Inventory (MBI) emphasizes three subscales, namely: emotional exhaustion, depersonalization, and reduced personal achievement. Most commonly used in measuring burnout in various professions. Job Demands-Resources (JD-R) Model summarizes the differences between the levels of job demands (workload, emotional demand) and the levels of job resources (support, decision making) whereby imbalance leads to burnout(Slemp & Vella-Brodrick, 2013). Conservation of Resources (COR) Theory (Hobfoll, 1989), proposes that burnout happens when people notice that their resources are diminished, or when threats are made to those resources, particularly after stress cycles without the ability to recuperate. Stress is a major issue, especially in professions that require individuals to be highly productive, contributing to worsening burnout, productivity, and employee well-being. Since organizations and researchers continue to work towards addressing burnout, predictive models play the role of supporting organizations in minimizing the likelihood of such issues. This literature review discusses the current developments on burnout prediction techniques with the use of machine learning models, statistical models, and psychometric models.

Even with the advancement of technology, there has been a rise in the rate of burnout among workers in organizations. Mental health in organizations has become priorities of employers as it pays attention to the psychological and emotional state of being a worker. Adaptation of a positive work culture promotes a good state of the organization's mental health in general, which reduces or avoids stressors like excessive overtime, job insecurities, heavy workloads, and tight deadlines. A study published in the Journal of Occupational and Environmental Medicine found that workers with higher levels of stress were more likely to report lower job satisfaction, contributing to higher turnover rates and absenteeism(Tatsuse et al., 2019). Psychological studies, questionnaires on burnout and its influencing factors have helped in creating theories and models to predict and avert a high rate of burnout (Shoji et al., 2016).

2.1.1 Traditional methods and approaches

Mindfulness and emotional regulation have been shown to significantly influence mental health outcomes among workers (Grégoire et al., 2015). The concept of the model of Effort-Reward Imbalance (ERI) proposed by Siegrist in 1996 states that burnout may arise from an imbalance between the incentives (such as compensation, recognition, or professional development) and the work put into a job. Perceived Stress Scale (PSS) and resilience evaluations can assist in measuring stress levels and resilience. Research on burnout has typically emphasized organizational and job-related factors, often explored within the Job Demands–Resources (JD-R) model (Bakker & Demerouti, 2018). In contrast, individual-level factors have received comparatively less attention. Nonetheless, several meta-analyses suggest personality plays a significant role in burnout (Alarcon et al., 2009; Kim et al., 2019). Specifically, personality traits from the Five-Factor Model (FFM) have been linked to burnout tendencies, Swider & Zimmerman (2010) found that high Neuroticism correlates with greater susceptibility to burnout, whereas traits such as extraversion, conscientiousness, and agreeableness are associated with lower burnout risk.

2.2 Concept of Machine Learning for Burnout Prediction

Machine learning is an enormous branch of artificial intelligence that entails training algorithms, which, with time, enhance their performance by identifying patterns and associations in data. Machine learning allows the model to learn how to make predictions from data (Holzinger, 2019). Machine learning has three main subdivisions, namely, supervised learning, unsupervised learning, and reinforcement learning. Models used for prediction include logistic regression, linear regression, support vector machines (SVM), decision tree classifiers, gradient boosting machines (like XG Boost), random forests, K-nearest neighbours (KNN), and neural networks. Under supervised learning, algorithms like linear regression find relationships between given input variables and output. Logistic regression is used for classification, usually binary classification. Supervised learning mainly predicts on labelled training data, while unsupervised learning does not. Different target variable types call for different approaches: regression is used for continuous data, and classification is used for categorical variables. (Sen et al., 2019). There are a few exceptions, such as K-Means Mode, which handles both kinds of data. Random forest models improve the interpretability of models through feature selection, making it easier to pinpoint specific job demands contributing to burnout. Gradient boosting models (GBM) such as XGBoost have been fine-tuned for predicting burnout by focusing on important variables like workload, support systems, and time at work. Support vector machine's (SVM) ability to define a boundary between burnout and non-burnout cases has been enhanced through hyperparameter tuning, making it a valuable model for early burnout detection, particularly when there's high-dimensional data (Ghavidel & Pazos, 2025). Refining these models, researchers and organizations can achieve more reliable burnout predictions, enabling earlier interventions and more tailored employee support. In this research paper and implementation, our baseline approach to achieving this project was with the Cross-Industry Standard Process for Data Mining (CRISP-DM), a widely-used framework for guiding data science projects. CRISP-DM provided clear advantages over earlier research in terms of model performance and process structure for the prediction project. CRISP-DM's structured six-stage strategy made sure that every phase was strategically in line with the project's objectives(Grządzielewska, 2021; Sen et al., 2019).

2.3 Machine learning in organizational psychology

Organizational psychology and machine learning are rapidly on the rise in organizational settings, especially for predicting and addressing psychological outcomes like engagement, satisfaction, turnover intentions, and well-being (Fukui & Wu, 2026). Through the analysis of big data: questionnaires and surveys of employees, indicators of productivity, communication behaviors, data from wearable devices, machine learning models can identify complex dependencies that are potentially biomarkers of risk or early manifestations of deterioration of psychological health (Răducu & Stănculescu, 2022; Shoji et al., 2016). For instance, the supervised learning models have been used in the past to predict the turnover risk using the job tenure, performance statistics, and frequency of feedback. In addition, state-of-the-art NLP methodologies enable the natural language analysis of the communications of the employees, including their e-mails or other feedback, to determine their emotions, stress, and satisfaction (Novikova, 2023). Machine learning also contributes to the tuning of mental and engagement interventions in that organizations are able to address psychological stressors as they occur, by feeding real-time data. This increases the accuracy of risk predictions and enables early intervention that addresses the psychological needs of individuals in a manner that optimizes organizational health instead of worsening adverse outcomes.

2.4 Literature gaps

The results of the meta-analysis for the study by Brewer were dependent on the quality and availability of data from previous studies. (Brewer & Shapard, 2004). The study was cross-sectional, making it impossible to determine causality or the long-term effects of age or experience on burnout. The study used semi-structured interviews to collect data, which may be subject to interviewer bias and limited reliability. A major gap in the literature is the limited generalizability of existing burnout prediction models. Most studies rely on small samples drawn from a single organization or occupation and provide little to no external validation across different populations or work contexts. A recent systematic review and meta-analysis of 22 burnout prediction studies reported a pooled area under the curve (AUC) of only 0.72, concluding that progress in the field is constrained by methodological weaknesses, heterogeneous samples, and the absence of external validation (Shi et al., 2025). These limitations are due to the lack of interpretability for non-technical stakeholders. Many burnout prediction approaches offer limited insight into how individual-level risk estimates are generated, making it difficult for human resource professionals and occupational health practitioners to justify or act upon the results. Transparency in prediction methods is particularly important in organizational settings, where decisions must be clearly explained and ethically defensible.

3. METHODOLOGY

In this research, we employ the CRISP-DM to predict employee burnout in organizations using machine learning techniques. CRISP-DM is a widely adopted, systematic approach for data mining that provides a structured workflow for building predictive models. This methodology consists of six phases: Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, and Deployment. Each phase is iterative, and insights gained in one phase may lead to revisiting earlier phases.

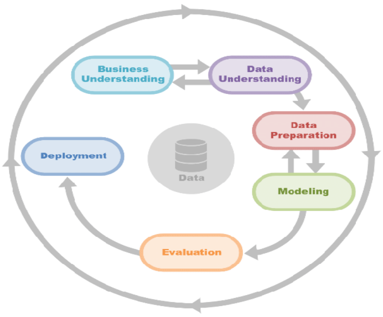

Figure 3.1: CRISP-DM Model (Tounsi et al., 2020)

3.1 Business Understanding

The first step in the CRISP-DM process is to understand the business objectives and define the problem. In this case, the goal is to predict employee burnout in an organization. Employee burnout is often characterized by emotional exhaustion, depersonalization, and a reduced sense of personal accomplishment. Predicting burnout can help organizations take proactive steps to mitigate its impact, improve employee well-being, and enhance productivity. The primary business question is: "How can we predict employee burnout based on historical employee data?"

3.2 Data Understanding

In this phase, we collect and explore data to gain insights into the factors contributing to employee burnout. The data used in this study comes from an employee survey that was carried out to source data for this research. The data was collected through Google Forms. It includes key variables such as:

1. **Demographic Data**: Age range.
2. **Job-related Data**: Workload, job satisfaction, role clarity, work-life balance, frequency of covering up for other employees.
3. **Psychological Factors:** Stress levels, emotional appreciation, and leave day satisfaction.
4. **Organizational Factors**: Team communication level, Team collaboration level, Working model, Work culture.

This data was organized and arranged using the key variables as represented in Table 3-1 below.

Table 3-1: Features Table

[1] **S/N**	[2] **Feature**	[3] **Type**	[4] **Description**
[5] 1	[6] Age Range	[7] Categorical	[8] Age group of the respondent (e.g., 16–25, 26–40).
[9] 2	[10] Work Model	[11] Categorical	[12] Type of work arrangement (Remote, Hybrid, On-site).
[13] 3	[14] Outstanding Tasks	[15] Numerical (Discrete)	[16] Number of pending work tasks at the time of response.
[17] 4	[18] Annual Leave Satisfaction	[19] Categorical (Ordinal)	[20] Level of satisfaction with allocated annual leave (Yes, No, Maybe).
[21] 5	[22] Feeling Appreciated	[23] Categorical (Binary)	[24] Indicates whether the respondent feels appreciated at work (Yes/No).
[25] 6	[26] Job Satisfaction	[27] Categorical (Ordinal)	[28] Indicates whether the respondent enjoys their job (Yes, No, Maybe).
[29] 7	[30] Work Culture Rating	[31] Numerical (Ordinal)	[32] Self-reported rating of workplace culture on a scale (e.g., 1–10).
[33] 8	[34] Manager Communication Rating	[35] Numerical (Ordinal)	[36] Rating of communication effectiveness with manager/employer/client (scale of 1–10).
[37] 9	[38] Team Collaboration Rating	[39] Numerical (Ordinal)	[40] Rating of collaboration within the team (scale of 1–10).
[41] 10	[42] Stress Level	[43] Numerical (Ordinal)	[44] Self-reported work-induced stress level (scale of 1–10).
[45] 11	[46] Covering for Colleagues	[47] Categorical (Ordinal)	[48] Frequency of covering for colleagues (Never, Seldom, Often, Very often).
[49] 12	[50] Task Overload	[51] Categorical (Binary)	[52] Indicates if the respondent frequently takes on more tasks than manageable (Yes/No).
[53] 13	[54] Burnout Status	[55] Categorical (Binary)	[56] Indicates whether the respondent is currently experiencing burnout (Yes/No).
[57] 14	[58] Comments	[59] Text	[60] Optional open-ended feedback provided by the respondent.

Data exploration often involves data visualization, data wrangling, and statistical data analysis. This step is crucial for understanding the underlying data and informing subsequent decisions regarding data preparation. Some of the Python modules used during this step are listed below.

1. **Pandas**: For data manipulation
2. **Matplotlib**: For data visualization
3. **Seaborn**: For advanced data visualization
4. **NumPy**: For numerical manipulations and calculations on arrays

Burnout Label Definition

The dataset uses self-reported burnout classification. This label ranges between two values: "Yes" or "No." The column had a distribution ratio of 16:11, with only 44 of the 108 respondents self-reporting burnout.

3.3 Data Preparation

Data preparation involves cleaning and transforming raw data into a suitable format for machine learning. This process involves:

1. **Data Cleaning**: In this step, we handle missing values, detect and correct outliers, and ensure uniformity in the dataset. Imputation techniques may also be applied for missing values and remove or adjust outliers that could bias the model.
2. **Feature Engineering**: This involves creating new features that change the representation of the underlying patterns in the data. For example, we could create composite scores for job satisfaction or emotional well-being.
3. **Data Encoding**: This involves converting categorical columns into numerical encodings or representations. This is crucial because a lot of ML models and algorithms only understand numeric input, but raw data usually has categorical (non-numeric) features, like text labels (for example, "Yes" or "No" responses or colors such as "Red," "Blue," and "Green"). This step was done using the Label Encoder function in scikit-learn.
4. **Normalization/Standardization:** Numerical features are scaled to ensure that all features are on a comparable scale and ensure optimal model performance. This is particularly helpful when working with those who rely on distance metrics such as K-Nearest Neighbors or Support Vector Classifiers.
5. **Data Splitting**: Data splitting, in this case, entails segmenting the data set into train and test segments in a manner that the machine learning algorithm is only trained on the train set and the test set is held out to test the accuracy of the trained model. In this study, the data were divided into train and test sets in a ratio of 80:20, and the test set was subsequently used to measure the model's performance on unknown data.

3.4 Modelling

In the modelling phase, we applied 7 different machine learning algorithms to train a predictive model for employee burnout. These algorithms were then tested and evaluated for their performance. They include:

1. **Logistic Regression**: This binary classification uses a logistic function to estimate the probability of class membership, a linear model. It performs nicely with linearly separable data and has interpretable coefficients for each feature.
2. **Support Vector Classifier**: This type of classifier separates classes by the hyperplane that covers them with the maximum margin between them. SVM is useful for high dimensions, and it performs well if the data can have clearly distinguished margins.
3. **Decision Tree Classifier**: This algorithm takes data and splits the data into branches depending on feature values to form the class tree. Decision trees are easy to interpret; however, the model tends to overfit without constraints.
4. **Random Forest Classifier**: This is an ensemble of decision trees that reduces overfitting by averaging the predictions of multiple trees but improves accuracy. Random Forest is robust and able to work well with both classification and regression tasks.
5. **K-Nearest Neighbour Classifier**: This is a model that classifies the data points according to the most common class of their nearest neighbors. KNN is simple and intuitive, and scales pretty well with large datasets. A disadvantage is that it can be slow.
6. **XGBoost Classifier**: This is a powerful gradient boosting model that builds trees sequentially and reduces errors made in the previous iterations. XGBoost is well known for its high performance and is widely used for competitive data science.
7. **Stacking Classifier (Random Forest + XGBoost + Logistic Regression):** This is an ensemble learning method, which involves using many base models for predictive performance enhancement. This Stacking Classifier uses Random Forest and XGBoost as the base learners to take into account various trends in the data, and a Logistic Regression model, which serves as a meta-learner, to arrive at the final decision. It can utilize the strong points of several models to lessen individual model partiality and enhance the overall accuracy and generalization.
8. **Feed-Forward Neural Network (FFN)**: A network of interconnected nodes with definitions of relationships between them, and a set of procedures that are applied to that network to learn complex patterns. FFN is very versatile and can model non-linear relationships, but it takes large datasets to work well.

3.5 Evaluation

Once the models were trained, they were then evaluated using the measurement metrics highlighted below. The results of the model evaluation were then used to select the optimal model to be used for deployment in the next stage. These metrics are:

1. **Accuracy Score:** This measures the correctness of the predicted instances out of the total instances.

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

Where;

TP=True Positives; TN=True Negatives; FP=False Positives; FN=False Negatives

2. **Precision**: True positives divided by the sum of true positives and false positives. It indicates how many of the predicted positives are actual positives.

$$Precision = \frac{TP}{TP + FP}$$

3. **Recall:** The quotient of true positives and the sum of true positives and false negatives. Recall shows how many of those actual positives were actually found, which is important if you miss any positives.

$$Recall = \frac{TP}{TP + FN}$$

4. **F1-Score**: This gives the harmonic mean between precision and recall. F1-score is useful when the dataset is imbalanced

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

5. **ROC/AUC Score**: This measures how much the model can distinguish between classes. A better performance for ranking positives against negatives is represented by a higher AUC.

$$AUC = \int_{FPR=0}^{1} \frac{TP}{TP + FN} . d(FPR)$$

Where; $FPR = \frac{FP}{FP+TN}$

6. **Loss**: A definition of a number (function) that measures the difference between predicted and actual, which leads to the optimization of the model. The generally higher values of loss indicate that models perform well.

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

where y_i is the true label, p_i is the predicted probability, and N is the number of samples.

3.6 Deployment

In this phase, the trained model is deployed into a real-world setting where it can be used to predict employee burnout in organizations. This research utilized Streamlit, an open-source Python package that assists developers in quickly creating interactive web applications for data science and machine learning projects. It is especially useful when deploying machine learning models because developers are able to create user interfaces with inputs, which can be provided as inputs, and model predictions or results can be shown in real time. Additionally, Streamlit comes with an automatic web framework setup that helps data scientists and machine learning engineers who may not have a solid frontend background. It also offers an easy way to manage model results, visuals, and user inputs to make it a popular tool for fast prototyping and demonstration of machine learning applications.

4. RESULTS AND DISCUSSIONS

Discussions

Descriptive analysis of the datasets in this study was conducted using the Python libraries seaborn and matplotlib. These tools were employed to support systematic data manipulation and visualization, enabling the identification of key patterns, trends, and relationships within the data. The resulting insights are presented and discussed below.

4.1 Stress Distribution among Employees

Employees experiencing burnout tend to have a wide range of stress levels, with a peak in the middle (around level 6-7), and some extending to higher stress levels (up to 10). In contrast, employees who are not burned out have a more concentrated stress distribution, peaking around level 6. Although some stress levels overlap, employees experiencing burnout tend to exhibit higher stress levels, which may reflect the additional mental strain and pressure associated with burnout.

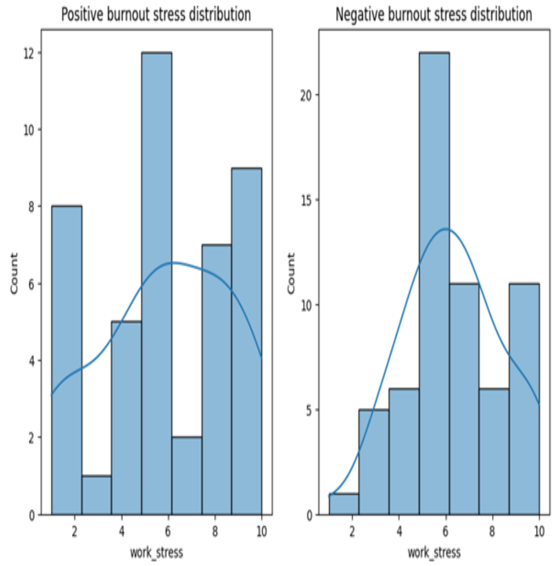

Figure 4.1: Stress Distribution

4.1.1 Annual Leave Satisfaction among Employees

Among employees experiencing negative burnout, a majority (51.6%) are satisfied with their annual leave days, while only 24.2% are dissatisfied. In contrast, only 20.5% of employees with positive burnout report satisfaction with their leave days, and nearly half (47.7%) express dissatisfaction. This implies that satisfaction with annual leave may play a role in preventing burnout. Employees who feel they have adequate time to recharge are less likely to experience burnout, while dissatisfaction with leave could indicate insufficient time off, leading to fatigue and a higher likelihood of burnout. Organizations could consider evaluating their leave policies and ensuring that employees feel they have sufficient time for rest.

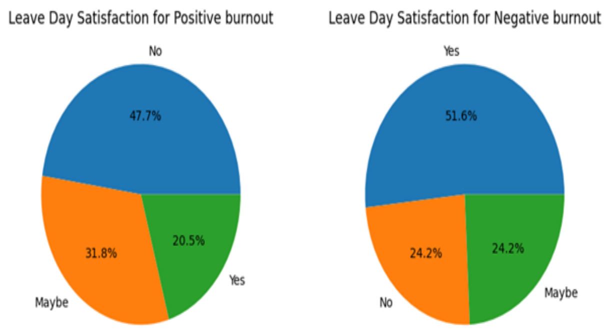

Figure 4.2: Annual Leave Satisfaction

4.1.2 Work Appreciation Influence on Burnout Rate

A large portion (62.9%) of employees with negative burnout feel appreciated at work, while only 8.1% do not feel appreciated. In contrast, a lower percentage (40.9%) of employees with positive burnout feel appreciated, and a significant 36.4% report not feeling appreciated. This suggests that feeling valued and acknowledged is an essential factor in mitigating burnout. Employees who do not feel appreciated may become disengaged and experience increased burnout levels. This data implies that organizations should actively recognize and appreciate employees' efforts, as a lack of appreciation could contribute to burnout.

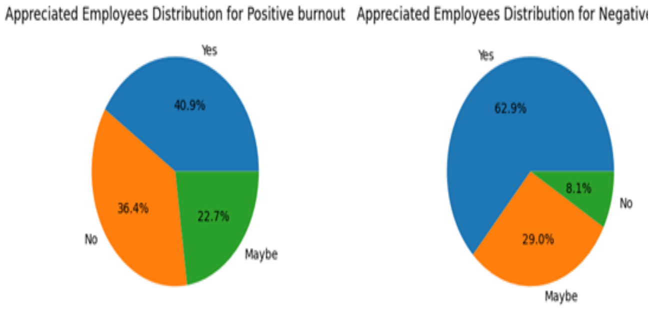

Figure 4.3: Work Appreciation

4.1.3 Frequency of Covering for Other Employees

In terms of covering for others, 62.9% of employees with negative burnout seldom cover up for others, 29% often do, and only 8.1% never cover for others. Conversely, among employees with positive burnout, a notable 20.5% very often cover for others, 18.2% often cover, and 54.5% seldom cover, while only 6.8% never do. This suggests that frequent covering for others may be linked to increased burnout, as employees may feel burdened by additional responsibilities beyond their regular duties. Organizations could examine how often employees are asked to cover for others and consider hiring additional support if necessary. Frequent task-shifting or added responsibilities can contribute to burnout, so managing these expectations may help prevent it.

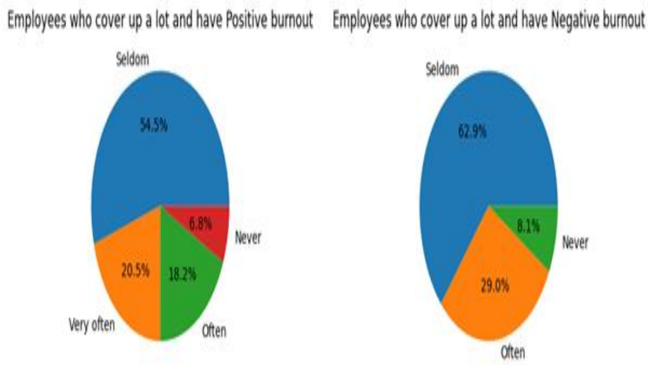

Figure 4.4: Frequency of covering up for other employees

4.1.4 Workload Influence on Burnout Rate

A notable 72.6% of employees with negative burnout are not overworked, while the remaining 27.4% are. In comparison, a significant 72.7% of employees with positive burnout feel overworked, with only 27.3% reporting a manageable workload. This finding shows the strong link between excessive workload and burnout. Overworked employees are more likely to experience burnout.

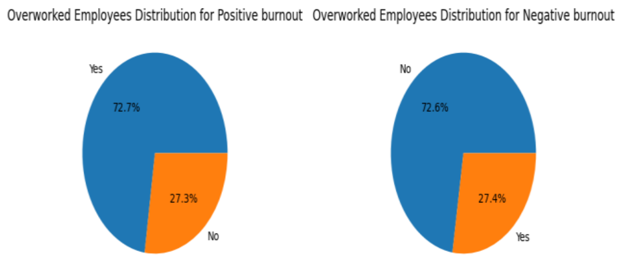

Figure 4.5: Workload Influence on Burnout Rate

4.1.5 Model Development and Evaluation

To test our proposed model, we used 20% of the total collected dataset to evaluate the performance of our model. The performance metrics used to evaluate the model are accuracy score, precision & recall, confusion matrix, and ROC/AUC score.

Table 4-1: Model Evaluation Table

[61] S/N	[62] Model name	[63] Accuracy score	[64] Loss	[65] Precision	[66] Recall	[67] F1-Score	[68] ROC/AUC score
[69] 1	[70] Logistic Regression	[71] 0.8	[72] -	[73] 0.80	[74] 0.80	[75] 0.80	[76] 0.90
[77] 2	[78] Decision Tree Classifier	[79] 0.975	[80] -	[81] 0.98	[82] 0.97	[83] 0.97	[84] 0.97
[85] 3	[86] Random Forest Classifier	[87] 0.975	[88] -	[89] 0.98	[90] 0.97	[91] 0.97	[92] 1.0
[93] 4	[94] Support Vector Classifier	[95] 0.9	[96] -	[97] 0.9	[98] 0.9	[99] 0.9	[100] 0.98
[101]	[102] KNeighbors Classifier	[103] 0.675	[104] -	[105] 0.69	[106] 0.68	[107] 0.67	[108] 0.85
[109]	[110] XGBoost Classifier	[111] 0.975	[112] -	[113] 0.98	[114] 0.97	[115] 0.97	[116] 1.0

[117]	[118] Stacking Classifier (Random Forest + XGBoost + Logistic Regression)	[119] 0.975	[120]	[121] 0.98	[122] 0.97	[123] 0.97	[124] 1.0
[125]	[126] Feed-forward Neural Network	[127] 0.925	[128] 0.1813	[129] -	[130] -	[131] -	[132] -

5. RESULTS

As illustrated in Table 4-1, the Logistic Regression model achieved an accuracy of 0.80 and a ROC/AUC score of 0.90, with precision, recall, and F1-score all at 0.80. This indicates a solid baseline performance for binary classification tasks, though it is outperformed by more complex models. The Decision Tree Classifier, Random Forest Classifier, and XGBoost Classifier all achieved high performance, each with an accuracy of 0.975, precision of 0.98, recall of 0.97, and F1-score of 0.97. Additionally, both Random Forest and XGBoost recorded a perfect ROC/AUC score of 1.0, while the Decision Tree achieved 0.97. Similarly, the Stacking Classifier (Random Forest + XGBoost + Logistic Regression) also achieved identical performance metrics, including an ROC/AUC score of 1.0.

The Support Vector Classifier (SVC) demonstrated strong performance with an accuracy of 0.90 and a ROC/AUC score of 0.98, alongside balanced precision, recall, and F1-score values of 0.90, indicating good generalization ability. In contrast, the K-Nearest Neighbors (KNN) model performed the weakest, with an accuracy of 0.675, a precision of 0.69, a recall of 0.68, an F1-score of 0.67, and an ROC/AUC score of 0.85, suggesting limited effectiveness for this task. The Feed-forward Neural Network achieved an accuracy of 0.925 with a loss of 0.1813. However, the absence of additional evaluation metrics such as precision, recall, and ROC/AUC limits a comprehensive assessment of its performance.

Overall, while multiple models achieved high predictive performance, the XGBoost Classifier and the Stacking model demonstrated the best combination of accuracy and ROC/AUC performance. The Stacking Classifier Model was later selected for deployment.

4.2.1 Deployment of Burnout Prediction Application

The Stacking Classifier model was deployed to a web application using Streamlit, popularly used for deploying data science and machine learning applications. The application utilized the trained model, which was stored as a serialized object (a pickle file). The application additionally made use of the encoder and scaler files that were used to encode the categorical columns and scale the entire dataset during the data preparation stage of the CRISP-DM model.

The application takes in 12 user input fields and returns a prediction that tells the user whether or not they are burnt out, includes the confidence interval of the model's prediction, and gives a brief recommendation on what they can do to combat burnout or to avoid burnout. This functionality aims to make the burnout prediction model more accessible to individuals and organizations, fostering proactive approaches to managing and preventing employee burnout. The web application can be accessed at https://employee-burnout-prediction.streamlit.app/.

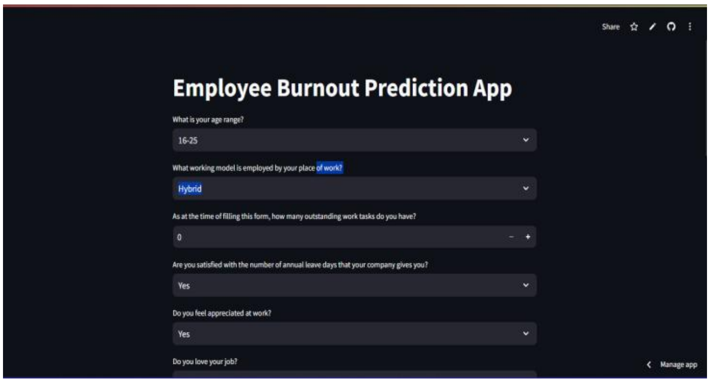

Figure 4.6: Deployed Web Application

6. Limitations of Results

The major limitation of the results in this chapter arises primarily from the small sample size of the dataset. While only 108 respondents were included in the data, they may not be good enough to represent the broader population, in which the model might not be generally applicable to other organizations or industries. Moreover, since the dataset was created from the answers in self-reported surveys, the data may be biased, by creating the possible bias in the response or social desirability bias. Furthermore, applying the model to a larger and more diverse dataset may result in drastically different performance of the model, making the collection and validation of more data, in diverse organizational contexts, necessary to validate the model's authenticity and reliability.

7. CONCLUSION

7.1 Summary of Key Findings

This research has shown that identifying some of the key factors that contribute to employee burnout can be predicted through the use of machine learning, namely, elements of the work environment, job satisfaction, and employees' personal perceptions of the work. During the data understanding stage, it was discovered that factors such as workload management, work culture, team collaboration & communication, and annual leave satisfaction are important factors in determining the likelihood of burnout. Through the development and evaluation of 7 different machine learning models, it was discovered that the XGBoost model had the most optimal metrics and was selected for deployment. These results suggest that machine learning tools may be useful for early detection of burnout and for timely intervention.

7.2 Implications for Organizations

This paper utilized a machine learning approach in predicting employee burnout and has practical benefits for organizations that want to promote a healthy workplace. Predictive models are leveraged to identify employees at greater risk of burnout, which then provides organizations with an outlet to enact proactive interventions. They can also give more aid to employees suffering from the initial signs of burnout, such as a flexible workload, as well as adequate rest. Applying this tool with a human-AI decision support system can boost employee well-being, productivity, and retention rates, cut absenteeism and poor performance at work, but most importantly, increase job satisfaction. Organizations should implement burnout prediction models in order to do a better job of anticipating workforce needs and proactively developing strategies that lead to a healthier, more engaged workforce.

7.3 Ethical Considerations

In this study, ethical standards were followed both in the collection and analysis of data. All participants received informed consent before filling out the survey, and the objective of the research and the intended use of the data were well communicated to them. It was voluntary, and respondents could pull out at any point with no repercussions. To ensure privacy, no personal identifiable information was collected, and all responses were fully anonymized. The dataset was used solely for academic research purposes, specifically for training and evaluating the machine learning model, and was not shared for commercial or external use.

Regarding ethical oversight, the study followed general research ethics guidelines. However, formal Institutional Review Board (IRB) approval was not required due to the non-sensitive and anonymous nature of the data collected.

7.4 Recommendations for Future Research

Future research should study larger, more diverse datasets that represent more diverse industries and roles in order to make machine learning models for predicting burnout robust and generalizable. Furthermore, the features of predictive modelling could be modified to include broader measures of work-life balance, mental health support, and roles of the job. Finally, the integration of qualitative data from employee feedback could also refine the models and give more personalized burnout.

8. REFERENCES

[1] Alarcon, G., Eschleman, K. J., & Bowling, N. A. (2009). Relationships between personality variables and burnout: A meta-analysis. *Work & Stress*, *23*(3), 244–263.

[2] Bakker, A. B., & Demerouti, E. (2018). Multiple levels in job demands-resources theory: Implications for employee well-being and performance. In *Handbook of well-being*. Noba Scholar.

[3] Brewer, E. W., & Shapard, L. (2004). Employee burnout: A meta-analysis of the relationship between age or years of experience. *Human Resource Development Review*, *3*(2), 102–123.

[4] Enzmann, D., Schaufeli, W. B., & Girault, N. (1995). The validity of the Maslach Burnout Inventory in three national samples. *Health Workers and AIDS: Research, Intervention and Current Issues in Burnout and Response*, 131–150.

[5] Fukui, S., & Wu, W. (2026). Prediction of Behavioral Health Employee Turnover With HR Data–Based Machine Learning Combined With Job Well-Being Indicators. *Psychiatric Services*, appi-ps.

[6] Gagnon, M.-C. J., Durand-Bush, N., & Young, B. W. (2016). Self-regulation capacity is linked to wellbeing and burnout in physicians and medical students: Implications for nurturing self-help skills. *International Journal of Wellbeing*, *6*(1), 101–116. https://doi.org/10.5502/ijw.v6i1.425

[7] Ghavidel, A., & Pazos, P. (2025). Machine learning (ML) techniques to predict breast cancer in imbalanced datasets: a systematic review. *Journal of Cancer Survivorship*, *19*(1), 270–294.

[8] Grégoire, S., Lachance, L., & Taylor, G. (2015). Mindfulness, mental health, and emotion regulation among workers. *International Journal of Wellbeing*, *5*(4).

[9] Grządzielewska, M. (2021). Using Machine Learning in Burnout Prediction: A Survey. *Child and Adolescent Social Work Journal*, *38*(2), 175–180. https://doi.org/10.1007/s10560-020-00733-w

[10] Hobfoll, S. E. (1989). Conservation of resources: a new attempt at conceptualizing stress. *American Psychologist*, *44*(3), 513.

[11] Holzinger, A. (2019). Introduction to machine learning & knowledge extraction (make). In *Machine learning and knowledge extraction* (Vol. 1, Number 1, pp. 1–20). Multidisciplinary Digital Publishing Institute.

[12] Huebner, L.-A., & Zacher, H. (2021). Following up on employee surveys: A conceptual framework and systematic review. *Frontiers in Psychology*, *12*, 801073.

[13] Kim, L. E., Jörg, V., & Klassen, R. M. (2019). A meta-analysis of the effects of teacher personality on teacher effectiveness and burnout. *Educational Psychology Review*, *31*(1), 163–195.

[14] Liu, Y., Cherian, J., Ahmad, N., Han, H., de Vicente-Lama, M., & Ariza-Montes, A. (2023). Internal corporate social responsibility and employee burnout: an employee management perspective from the healthcare sector. *Psychology Research and Behavior Management*, 283–302.

[15] Liu, Z., Zhou, Y., Xu, Y., & Wang, Z. (2023). Simplenet: A simple network for image anomaly detection and localization. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 20402–20411.

[16] Maslach, C. (2013). Burnout research in the social services: A critique. In *Burnout among social workers* (pp. 95–105). Routledge.

[17] Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology*, *52*(2001), 397–422.

[18] McCarthy, S. C. (2025). *The Relationship Between Self-Care, Resilience and Burnout, and the Role of Self-Esteem and Self-Awareness to Facilitate Self-Care Among School Counselors*.

[19] Novikova, E. (2023). *Influence of NLP on organizational design and development of dynamic capabilities*.

[20] Răducu, C.-M., & Stănculescu, E. (2022). Teachers' burnout risk during the COVID-19 pandemic: relationships with socio-contextual stress—a latent profile analysis. *Frontiers in Psychiatry*, *13*, 870098.

[21] Sen, P. C., Hajra, M., & Ghosh, M. (2019). Supervised classification algorithms in machine learning: A survey and review. In *Emerging technology in modelling and graphics: Proceedings of IEM graph 2018* (pp. 99–111). Springer.

[22] Shi, H., Liu, J., Yang, C., Shang, J., & Zeng, Y. (2025). Machine learning for predicting burnout among healthcare workers: a systematic review and meta-analysis. In *Contemporary Nurse*. Routledge. https://doi.org/10.1080/10376178.2025.2593294

[23] Shirom, A., & Melamed, S. (2006). A comparison of the construct validity of two burnout measures in two groups of professionals. *International Journal of Stress Management*, *13*(2), 176.

[24] Shoji, K., Cieslak, R., Smoktunowicz, E., Rogala, A., Benight, C. C., & Luszczynska, A. (2016). Associations between job burnout and self-efficacy: A meta-analysis. *Anxiety, Stress, & Coping*, *29*(4), 367–386.

[25] Slemp, G. R., & Vella-Brodrick, D. A. (2013). *The Job Crafting Questionnaire: A new scale to measure the extent to which employees engage in job crafting.* Https://Internationaljournalofwellbeing.Org/Index.Php/Ijow/Article/View/217.

[26] Spector, P. E. (2022). *Job satisfaction: From assessment to intervention*. Routledge.

[27] Swider, B. W., & Zimmerman, R. D. (2010). Born to burnout: A meta-analytic path model of personality, job burnout, and work outcomes. *Journal of Vocational Behavior*, *76*(3), 487–506.

[28] Tatsuse, T., Sekine, M., & Yamada, M. (2019). The Contributions Made by Job Satisfaction and Psychosocial Stress to the Development and Persistence of Depressive Symptoms. *Journal of Occupational & Environmental Medicine*, *61*(3), 190–196. https://doi.org/10.1097/JOM.0000000000001491

[29] Tehrani, F. S., Calvello, M., Liu, Z., Zhang, L., & Lacasse, S. (2022). Machine learning and landslide studies: recent advances and applications. *Natural Hazards*, *114*(2), 1197–1245.

[30] Wahab, A., & Anwar, M. (2023). The effect of work placement and work discipline on employee morale through organizational commitment as an intervening variable at the Mercury hotel banjarmasin with the SEMPLS-Warppls 7.0 application. *Int. J. Soc. Sci. Educ. Res. Stud*, *3*, 1146–1148.

5. CODE AND DATA AVAILABILITY

To support reproducibility and further research, the complete implementation of the proposed system is publicly available. The source code, including the model training file, the web API code file, and the web application interface, can be accessed via the GitHub repository: [https://github.com/Nalito/employee-burnout-pred].

A live version of the application for real-time burnout inference is also available at: [https://employee-burnout-prediction.streamlit.app/].

Due to limitations in publicly available datasets for this task, a custom dataset was constructed and used for training. Details on data collection, preprocessing, and usage guidelines are provided within the repository. Instructions for environment setup, dependencies, and reproduction of results are included in the project documentation to ensure ease of replication.