Enhancing worker well-being by utilising Data Analytics and Machine Learning approaches for fatigue detection

MSc Research Project
Data Analytics

Anish Romario Joseph Ambrose
Student ID: x20190841

School of Computing
National College of Ireland

Supervisor: Abdul Qayum

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Anish Romario Joseph Ambrose
Student ID:	x20190841
Programme:	Data Analytics
Year:	2023
Module:	MSc Research Project
Supervisor:	Abdul Qayum
Submission Due Date:	14/12/2023
Project Title:	Enhancing worker well-being by utilising Data Analytics and Machine Learning approaches for fatigue detection
Word Count:	
Page Count:	14

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	
Date:	31st January 2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Enhancing worker well-being by utilising Data Analytics and Machine Learning approaches for fatigue detection

Anish Romario Joseph Ambrose
x20190841

Abstract

This research addresses the imperative of employee well-being through the exploration and development of machine learning and data analytics for early detection of fatigue in diverse workplace environments. Utilizing models such as Decision Tree, Feed-forward Neural Network, Deep Learning, K-Nearest Neighbours (KNN), XG Boost, and Random Forest, we aim to identify early signs of fatigue and stress across various industries.

Our primary objectives include the creation of predictive models capable of analysing multiple data sources to discern patterns associated with fatigue. Noteworthy achievements include the Feed-forward Neural Network and Deep Learning algorithms portraying superior predictive capabilities with low Mean Squared Error values and High R-squared values. Moreover, this study assesses the broader impact of data-driven fatigue detection systems on workplace safety, employee well-being, job performance, and job satisfaction.

The findings emphasize the efficiency of the Random Forest model in promoting workplace safety. By addressing these objectives, this research contributes valuable insights to the development of proactive strategies for detecting and mitigating employee fatigue, ultimately fostering healthier and more productive work environments.

1 INTRODUCTION

In the realm of employee well-being, particularly within the dynamic and demanding context of the technical industry, the need for effective fatigue detection systems becomes paramount. This research embarks on a journey to design and implement data analytics and machine learning models that discern early signs of fatigue among workers, with the ultimate goal of enhancing workplace safety and employee welfare. The motivations behind this project stem from the intricate challenges posed by the modern work landscape, marked by globalization, technological advancements, and heightened performance expectations [Hooda et al. (2021)] [Parekh et al. (2020)].

1.1 Background to the Problem

The technical industry, among others, faces complexities and unpredictability's, necessitating a closer examination of factors influencing employee well-being. Fatigue, often arising from prolonged work hours, insufficient rest, and the demanding nature of technical projects, emerges as a critical concern. This research delves into the nuanced aspects of fatigue detection, aiming to bridge a major gap in the current literature by developing accurate, efficient and scalable algorithms tailored for the technical sector [Hooda et al. (2021)] [Parekh et al. (2020)].

1.2 Motivations and Project Choice

The motivation behind choosing this project lies in the possibility of enhancing the well-being of employees in high-risk industries. By developing proactive fatigue detection systems, our goal is to lessen the possibility of accidents, enhance job performance, and establish a robust framework for occupational well-being. This research is not only academically significant but also holds industrial relevance by addressing a pressing concern in workplace safety [Hooda et al. (2021)] [Parekh et al. (2020)].

1.3 Research Question and Objectives

The central question driving this research is: How can machine learning, and data analytics models be effectively designed and implemented to detect early signs of fatigue in the technical industry? To address this, our objectives are threefold:

Investigate the use of machine learning and data analytics in identifying early signs of fatigue and stress. Develop predictive models capable of analysing diverse data sources to detect patterns associated with fatigue [Hooda et al. (2021)] [Parekh et al. (2020)].

Assess the impact of data-driven fatigue detection systems on workplace safety and employee well-being, as well as their potential to enhance job performance and satisfaction.

1.4 Methods

The methodology involves a comprehensive examination of various machine learning models, including Decision Tree, Feed-forward Neural Network, Deep Learning, K-Nearest Neighbors (KNN), XGBoost, and Random Forest. This research leverages a diverse dataset comprising physiological and cognitive features as the foundation for developing predictive models. The challenge lies in creating algorithms that are accurate, efficient, and scalable to meet the unique demands in the technical sector [Hooda et al. (2021)] [Parekh et al. (2020)].

1.5 Structure of the Report

This research unfolds in a structured manner to offer a cohesive narrative. Later the introduction and subsequent parts will examine into the methodology, data sources, and the use of data analytics and machine learning in addressing fatigue in diverse workplace settings. This study will also discuss potential challenges and ethical considerations associated with implementing these systems. In conclusion, we will summarize the anticipated impact on employee well-being and safety, resulting in an understanding of how data analytics and machine learning can play a significant role with respect to health and safety in the workplace.

2 LITERATURE SURVEY

Rohit Hooda and co-authors Hooda et al. (2021) delve into the ever-evolving landscape of technological progress, significantly impacting daily lives and introducing a new challenge—fatigue. The study explores fatigue's intricate facets, including manifestations, consequences, and the application of machine learning (ML) approaches for its detection. The paper categorizes methods into Mathematical Models, Machine Learning models, Rule-Based Implementation, and Deep Learning, providing a comprehensive overview of recent innovations in fatigue detection. The study meticulously compares various algorithms with the aim of identifying the most promising approach for fatigue detection.

Parekh and team Parekh et al. (2020) focus on the impact of technological advances in healthcare on patient outcomes and quality of life. Their review centers on the health indicator of fatigue, examining its close relationship with cognitive performance and health outcomes. The paper presents studies that enhance the understanding of fatigue, offering systematic approaches and detection methods. Artificial intelligence, particularly artificial neural networks, wavelet transform, and data analysis of various parameters, proves essential in monitoring and detecting fatigue.

Meng-Long Huo Ma et al. (2023) This literature review explores the nexus between employee well-being and lean production in a Chinese manufacturing context, employing the job demands–resources model. Examining problem-solving demands and job resources (training, decision-making participation, and line manager support), the study reveals a ”buffering effect,” mitigating the impact of demands on exhaustion, and a ”coping effect,” strengthening resource-engagement ties. Contrary to a uniform impact, the findings underscore the nuanced influence of lean production on worker well-being, contingent upon managerial practices in fostering involvement, aligning resources, and adjusting levels in response to job demands.

Tuncer and co-authors Tuncer et al. (2021) address the significant contribution of driver fatigue to traffic accidents, emphasizing the need for effective detection systems. Their study explores automated driver fatigue detection using electroencephalogram (EEG) signals. Benchmark classifiers demonstrate the efficacy of the proposed method, achieving a notable 97.29% classification accuracy for fatigue detection.

Choi and team Choi et al. (2019) investigate the use of physiological data, specifically Electrodermal Response (EDR) and Electrodermal Level (EDL), obtained from readily available wristband sensors to understand how technical workers perceive risk during their work. The research collected 30 hours of physiological data from eight technical workers as they carried out their usual tasks.

S. Ansari and colleagues Ansari et al. (2021) propose an approach to measure driver mental fatigue, analysing head angular acceleration data through a modified deep neural network and a long-term bidirectional memory. The classifier demonstrates superior performance, outperforming both conventional and machine learning techniques.

Shahzeb Ansari and colleagues Ansari et al. (2022) explore automatic detection of underload driver cognitive fatigue through upper body posture dynamics. Their study employs a semi-supervised approach with unsupervised Gaussian Mixture Model clustering. Machine learning classifiers achieve accurate recognition of various driving postures.

Joonchul Shin and colleagues Shin et al. (2019) employ the Smart Fatigue Phone to measure salivary cortisol concentration during a driving session. The system demonstrates a high correlation between alpha waves and cortisol concentration, offering a promising tool for real-time driver fatigue monitoring.

G. N, S. S, and colleagues Guk et al. (2019) propose a tool for continuous monitoring of facial expressions and landmarks to assess driver fatigue or emotional changes. Utilizing machine learning, particularly Support Vector Machines (SVMs), the developed system employs facial expressions to detect fatigue, triggering immediate alerts for proactive intervention.

Y. Li, D. Wang and co-authors Li et al. (2022) introduce the Auto-Correlation Function-based Sparse Support Matrix Machine (ACF-SSMM) algorithm, optimizing and classifying

EEG fatigue signals. The ACF-SSMM algorithm proves effective in EEG-based fatigue detection through improved results on the SEED-VIG dataset.

Zargari Marandi and colleagues Marandi et al. (2019) investigate the utilization of biofeedback techniques for real-time fatigue detection. Their study introduces a novel biofeedback system that combines physiological signals with user-reported data to provide personalized fatigue assessments. The research emphasizes the potential of biofeedback in creating adaptive interventions for fatigue management.

Alexis D. Souchet and co-authors Souchet et al. (2023) explore the application of virtual reality (VR) technology for fatigue assessment. Their study examines the use of immersive VR environments to simulate workplace scenarios, enabling a more ecologically valid evaluation of fatigue levels. The findings shed light on the integration of VR as a promising tool for stress-induced fatigue detection.

Chinoy, E.D., Cuellar, J.A., Jameson, J.T, and colleagues Chinoy et al. (2022) focus on circadian rhythm synchronization for fatigue mitigation in the context of shift work. Their study investigates the impact of aligning work schedules with individual circadian rhythms to reduce the risk of fatigue-related incidents. The research provides insights into chronobiological interventions for optimizing work schedules and promoting employee well-being.

Liu, P., Chi, H.L., Li, X, and colleagues Liu et al. (2021) contribute to the literature by exploring the role of dietary factors in fatigue management. Their study investigates the effects of nutritional interventions on energy levels and cognitive performance, underscoring the importance of dietary considerations in comprehensive fatigue prevention strategies.

Cui, Y., Zhang, M., and colleagues Cui et al. (2019) examine the intersection of sleep quality and fatigue in their study on wearable sleep tracking devices. The research evaluates the effectiveness of commercially available wearables in accurately assessing sleep patterns and its correlation with fatigue levels. The findings provide valuable insights into leveraging consumer-grade technology for fatigue monitoring.

Khosro Sadeghniiat and colleagues Oliver et al. (2021) contribute to the literature with their research on the widespread issue of fatigue among workers in today's industries, stemming from job pressures, long hours, sleep disruptions, and accumulated deficits. The severity of fatigue varies by industry, necessitating tailored approaches.

Santhosh, S, and colleagues Santhosh and Ar (2022) contribute to the literature with their research on mindfulness-based interventions for fatigue reduction. The study introduces a mindfulness training program aimed at enhancing attentional control and stress resilience. The findings highlight the potential of mindfulness practices in mitigating the cognitive and emotional aspects of fatigue.

3 RESEARCH METHODOLOGY

Providing transparency about the data collection methods is crucial for ensuring the credibility and reliability of your research. This research has employed a diverse set of machines learning models, ranging from traditional regression models to more advanced ensemble methods and neural networks. The features used for prediction include MEAN, MAX, MIN, RANGE, KURT, and SKEW. The models were trained and evaluated using ML and DL parameters such as MSE, R2, and classification metrics, depending on the nature of the model.

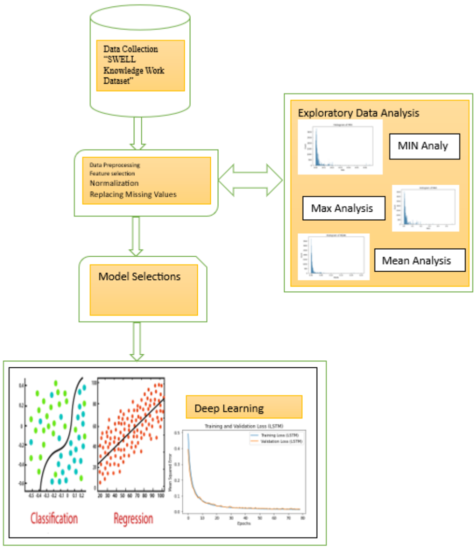

Figure 1: Methodology Flow Chart

3.1 Data Collection:

To address the objectives of the study, This dataset consists information on physiological measures, work hours, environmental conditions, and employee self-reported fatigue levels. Data sources may include wearables, sensors, employee surveys, and workplace records. This approach ensures a holistic representation of workplace conditions and allows the model to learn patterns from multiple dimensions, contributing to a robust fatigue detection system.

3.2 Data Preprocessing:

The Data preprocessing steps deals with outliers, normalize the numerical features, encode the categorical variables and also manage the missing values. Feature engineering will be employed to draw out suitable information from the original data and potentially incorporating time-series analysis for temporal patterns.[1]

3.3 Exploratory Data Analysis (EDA :

EDA involves visualizing and analysing the relationships within the dataset. This step aims to identify potential correlations between different features and fatigue levels. Insights gained from EDA will guide the selection of features for model development.

Justification: Understanding the data through EDA is crucial for informed feature selection, ensuring that the model focuses on relevant predictors of fatigue. This step aligns with the objective of exploring data analytics in identifying early signs of fatigue.

Figure 2 represents a histogram which is a pictorial depiction of a dataset's distribution. In the context we provided, the x-axis and y-axis respectively ranges from 0.00 to 0.25 and from 0 to 3500. The x-axis is divided into intervals or "bins" that cover the range from 0.00 to 0.25.

[1]http://cs.ru.nl/~skoldijk/SWELL-KW/Dataset.html [30]
Justification: For the purpose of training machine learning models data must be clear and organised. Feature engineering enhances the ability of the model to catch nuanced relationships between variables, contributing to the early detection of fatigue.

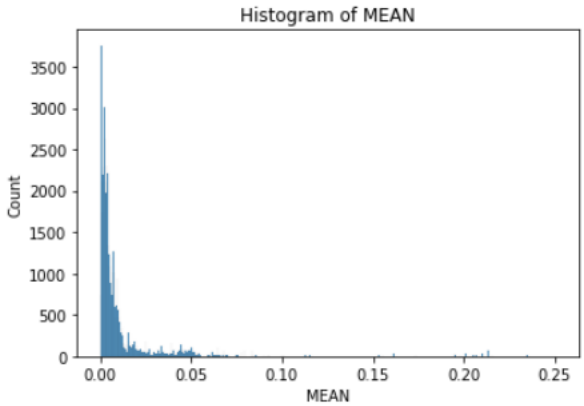

Figure 2: Histogram of MEAN

Each interval portrays a sub-range of values within the entire range. The frequency of data points appearing inside that interval is represented by the height of each bar in the histogram. Looking at the diagram. The bars on the histogram show how many data points or observations fall within each interval on the x-axis. The histogram's picture gives a detailed information on how the data are distributed. A symmetric distribution might suggest a normal distribution, while skewed distributions (positively or negatively) indicate asymmetry. Clusters or gaps in the data can also be identified Hooda et al.(2021) and Parekh et al.(2020).

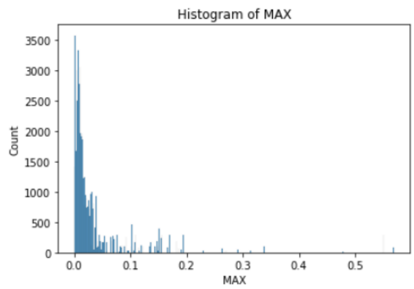

Figure 3: Histogram of MAX

A histogram of the variable "MAX," it means you're visualizing the distribution of maximum values within your dataset. A histogram of the variable "MAX" will show how the maximum values are distributed across different ranges. If there's a tall bar in a specific bin, it indicates that many data points have maximum values within that range Tuncer et al.(2021)

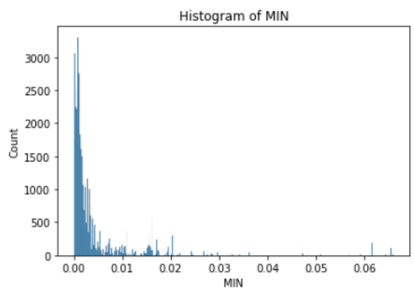

Figure 4: Histogram of Min

Figure 4 represents the histogram which visually represents the frequency distribution of minimum values, with a notable concentration at 0 and decreasing frequency as the minimum values deviate from this point. Analysing the histogram's form can offer good information with respect to the dataset's properties and guide further analysis or exploration. The dataset's minimal value distribution is revealed by the histogram and the concentration of data points at 0 suggests a specific pattern or characteristic in the data where many observations share a minimum value of 0. The spread and distribution beyond 0.02 and toward 0.06 provide information about the variability and outliers in the dataset Choi et al.(2019)

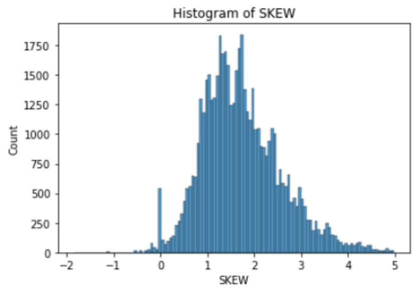

Figure 5: SKEW histogram

Figure 5 portrays the skewness values near 0 to be distributed symmetrically. As you move towards positive values, it suggests a rightward skewness (longer tail on the right), and as you move towards negative values, it suggests a leftward skewness (longer tail on the left). The distribution appears to be more concentrated within a certain range of skewness values. The frequency of observations is highest within this range. The histogram provides insights into the skewness of your dataset. A peak in frequency around a specific skewness value may indicate a predominant skewness direction in the data Ansari et al. (2021) and Ansari et al. (2022). [2]

3.4 Model Development:

Multiple machine learning models such as Random Forest, Deep Learning and Decision Tree were used to develop predictive models for fatigue detection. The models will be trained, tested and validated using a separate set to assess generalization performance. Different algorithms offer varying strengths in capturing complex patterns. Employing a variety of models enables the identification of the most suitable approach for early fatigue detection, aligning with the study's objective to develop predictive models.

3.4.1 Random Forest.

This model is selected for the thesis due to its ability to handle complex relationships within diverse workplace datasets. Its ensemble nature means combining weak decision tress to combining multiple decision trees, enables robust feature importance assessment and generalization to unseen data and accurately identifying the early signs of fatigue across different industries. A group of decision trees is constructed using Random Forest during training where each tree is built on a random subset of the characteristics and data. This randomness contributes to the model's diversity and reduces overfitting.

[2] https://www.investopedia.com/terms/s/skewness.asp#:~:text=Skewness%20is%20a%20measurement%20 of,median%20on%20a%20bell%20curve [19].

3.4.2 Decision Tree.

The model in this thesis is driven by its interpretability, ease of visualization, and capacity to capture non-linear relationships in complex datasets. Given the diverse workplace environments and the need for an interpretable model to understand the factors contributing to fatigue, a Decision Tree aligns with the study's objectives.

3.4.3 K-Nearest Neighbours (KNN):

This model is chosen for its simplicity and effectiveness in understanding the regional trends in the data. The ability of the model is to make predictions, based on the majority of class using its k-nearest neighbours is well-suited for identifying subtle patterns in the context of fatigue detection.

3.4.4 FNNs

are versatile and can capture intricate patterns in data making them suitable for regression tasks and their ability to model non-linear relationships is crucial in predicting continuous outcomes, such as fatigue levels in the context of the presented thesis.

3.4.5 Using LSTM

for regression in the context of fatigue detection aligns with the nature of the data, which often involves time-dependent patterns. These type of models are capable of recording enduring dependencies and consecutive data making them well-suited for tasks where the temporal aspect is crucial. Additionally, the adaptability of LSTMs to various input features and the ability to learn from complex temporal relationships make them a suitable choice for regression in dynamic and evolving workplace environments.

3.4.6 Gradient Boosting

is a powerful machine learning technique and this model help us to build predictive models in a sequential manner by combining the predictions of poor learners' which are usually decision trees of the model DT. The primary goal of gradient boosting is to reduce the residuals of the previous models, leading to a strong predictive model. The term "gradient" refers to the optimization method used to minimize the loss function.

3.4.7 XG Boost

which builds an ensemble of poor learners (typically decision trees) and consecutively enhances their performance which is well known for being accurate and making it a preferred choice in various data science competitions and real-world applications. The algorithm can handle missing values in the dataset which can automatically handle them during the training process.

3.5 Evaluation Metrics:

The proper measures such as precision, accuracy ,recall, and F1-score, Square root and Mean Square Error will be used to assess the model's performance. The choice of metrics will be driven by the balance needed between false positives and false negatives in fatigue detection. The selected metrics directly relate to the study's goal of developing accurate predictive models.

Precision and recall are especially important for assessing the impact of false positives, true positive and false negatives on workplace safety and employee well-being.

By utilizing a paradigm that reflects the complexity of the technical sector and delving into Deep Learning, to aid in the development of a reliable and precise fatigue detection system. Fig 6 and Fig 7 depicts the training and validation loss of deep learning model (LSTM) and feedforward Neural Network.

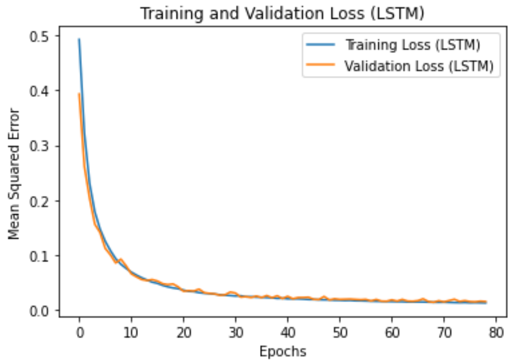

Figure 6: Training loss and Validation loss of LSTM

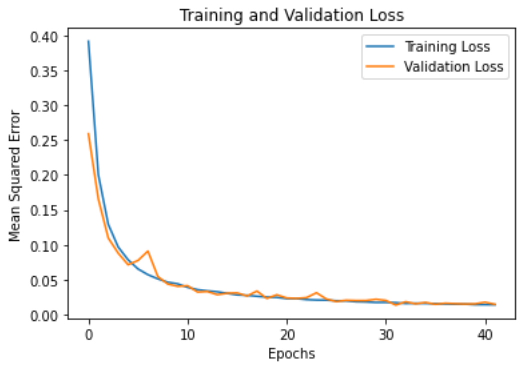

Figure 7: Training loss and Validation loss of feedforward Neural Network

3.6 CASE STUDY

A number of case studies have illustrated the potential of data analytics and machine learning models in augmenting employee well-being through identifying the indicators of fatigue. Proactive detection of fatigue indicators among technical workers was the objective of the development of predictive algorithms by industry researchers. These algorithms operated by analysing data derived from occupational sensors and cognitive processes. Accidents occurred to a significant lesser extent as a result of this strategy, which also increased employee satisfaction. The tension levels of healthcare personnel may be effectively detectable through continuous monitoring with ubiquitous technology, according to researchers. Consequently, this facilitates the creation of individualized solutions and a general enhancement in occupational contentment. Furthermore, transportation companies observed a substantial reduction in incidents associated with fatigue subsequent to the integration of fatigue detection algorithms into the systems of their vehicles. Through exemplifying the adaptability of machine learning and data analytics in a variety of professional settings, these case studies draw attention to the potential for enhancing employee productivity and wellness in the workplace.

4 Design Specification

Design and implement a Fatigue Detection System using a combination of Machine Learning and Deep Learning models for regression and classification associated preprocessing techniques to identify early signs of fatigue in workplace environments.

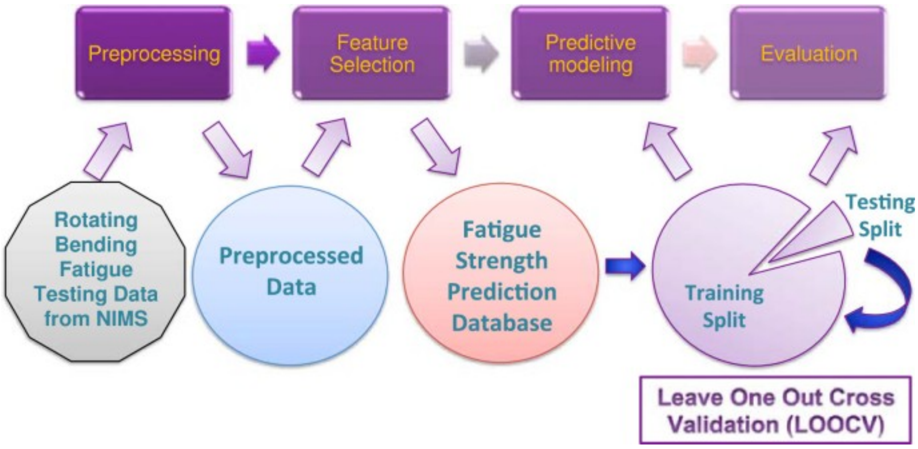

Figure 8: Exploration of data science techniques to predict fatigue (curtesy: Google)

Machine Learning for Classification:

Utilizing classic machine learning models such as Random Forest, XG Boost, Decision Tree and K-Nearest Neighbours (KNN) for classification drawing insights from [Hooda et al. (2021)] [Parekh et al. (2020)]. Select relevant features based on exploratory data analysis (EDA) [Hooda et al. (2021)]. Implementation of hyperparameter tuning to optimize the performance of classification models was obtained from [Parekh et al. (2020)]..

Deep Learning for Regression:

To capture sequential patterns in time-series data related to fatigue use neural networks equipped with Long Short-Term Memory (LSTM) [Ansari et al. (2021)]. To stop the deep learning model from overfitting use dropout layers [Ansari et al. (2021)].

Data Preprocessing Techniques:

Organization of time-series data into sequences for LSTM input in the regression model [Ansari et al. (2021)]. Numerical features are normalized for standard scaling in both regression and classification [Shin et al. (2019)].

Handle missing values and outliers appropriately.

Encode categorical variables for machine learning classification models and divide the dataset into training, validation, and test sets for both regression and classification tasks [Li et al. (2022)].

Evaluation Metrics:

These results provide insights from the referred data into the performance of each model in detecting fatigue. The key metrics such as Mean Squared Error, R-Squared, F1-score , Precision, Accuracy and Recall are used to evaluate their effectiveness. The lower the MSE and the higher the R-Squared, Precision, Recall, F1-score, and Accuracy, the better the model's performance [Marandi et al. (2019)] [Souchet et al. (2023)].

Among the models, the models such as Feed Forward Neural Network and Deep Learning exhibit low MSE and high R-Squared, indicating their strong predictive capabilities [Chinoy et al. (2022)]. XG Boost also demonstrates excellent precision, F1-score ,recall and accuracy. These results can guide in choosing the most suitable model for fatigue detection based on the specific requirements and priorities [Liu et al. (2021)].

5 RESULTS

In evaluating various machine learning models for a specific task, several key performance metrics were considered.From the Table 1, The decision tree model exhibited a Mean Squared Error (MSE) and R-Squared value of 33.831704 and 0.843, and an accuracy of 96%. On the other hand, the models such as Feed-forward Neural Network and Deep Learning outperformed with significantly lower MSE values such as 0.010 and 0.012, respectively, and impressive R-Squared values such as 0.981 and 0.983. The K-Nearest Neighbours (KNN) model demonstrated good precision (0.88), recall (0.87), F1-score (0.87), and an accuracy of 86%. XG Boost and Random Forest model demonstrated excellent precision ,recall and F1-score with values such as 0.96, 0.95, 0.95 and an accuracy of 96%. In summary, the neural network models, particularly the Deep Learning model, along with XG Boost and Random Forest, emerged as strong performers.

Table 1: Performance metrics of each model Classification

Decision Tree Classifier	Precision	0.95
	Recall	0.93
	F1-score	0.94
	Accuracy	0.96
Feed forward Neural Network	MSE	0.010
	R-squared	0.983
Deep Learning	MSE	0.012
	R-squared	0.981
KNN	Precision	0.88
	Recall	0.87
	F1-score	0.87
	Accuracy	0.86
XG Boost	Precision	0.96
	Recall	0.95
	F1-score	0.95
	Accuracy	0.96
Random Forest	Precision	0.96
	Recall	0.95
	F1-score	0.95
	Accuracy	0.96

Table 2: Regression model results

Model	**Mean Squared Error**	**R-square**
Random Forest	27.847374	0.871411
Decision Tree	33.831704	0.843778
Gradient Boosting	35.111897	0.837866
Linear Regression	213.684578	0.003974
Optimised Decision Tree	33.134627	0.846997

From the Table 2 Based on the provided information, the Random Forest seems to be the best performer for the prediction of fatigue. This conclusion is based on both the Mean

Squared Error and R-squared and this model appears to be the perfect choice from the provided key(evaluation) metrics.

Table 3: Temporal Analysis

Time Period	Fatigue Level (%)
Morning	22
Afternoon	29
Night	32

Table 3 represents a temporal analysis of fatigue levels, measured in percentages, across different time periods. In the morning, the fatigue level is recorded at 22%, indicating a relatively lower level of fatigue during this period. Moving to the afternoon, there is a slight increase in fatigue, with the level reaching 29%. The highest fatigue level is observed at night, with a recorded percentage of 32%, suggesting a potentially increased sense of fatigue during nighttime. This temporal breakdowngives insightful observations into the fluctuation of fatigue levels throughout the day which can be crucial for understanding patterns and making informed decisions related to factors influencing fatigue management.

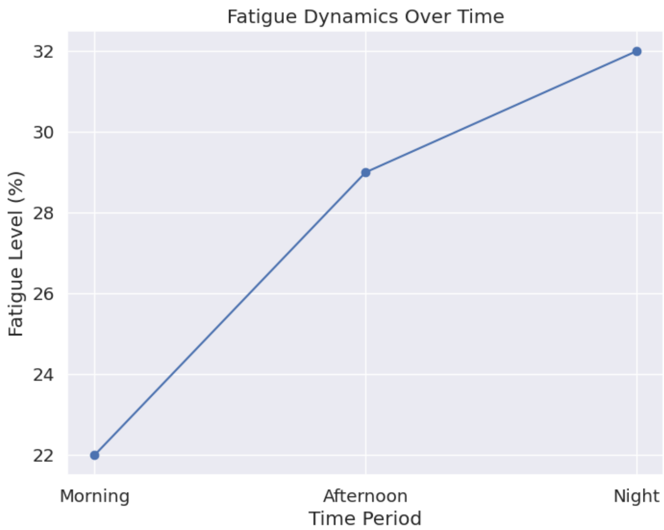

Figure 9: Fatigue dynamics over time (curtesy: Google)

6 CONCLUSION

In conclusion, the comprehensive evaluation of other machine learning models for the specific task yielded valuable insights into their performance metrics. Notably, the models such as feed-forward neural network and deep learning demonstrated exceptional results, boasting low Mean Squared Error values such as 0.010 and 0.012, respectively, and high R-Squared values such as 0.983 and 0.981. XG Boost and Random Forest also portrayed good performance as measured by recall, precision, accuracy and F1-score. These findings underscore the effectiveness of neural network and ensemble methods in the given task, providing a basis for informed model selection.

7 FUTURE WORK:

Moving forward, there are many other avenues for future work. Firstly, lets go more in-depth exploration of feature engineering could be undertaken to improve the model's capabilities for prediction. Furthermore, investigating the effects of additional data sources or refining existing datasets may contribute to better model generalization. Fine-tuning hyperparameters and exploring advanced neural network architectures could further optimize model performance. Moreover, conducting a more granular temporal analysis, considering factors such as day of the week or specific hours within each time period, could unveil nuanced patterns in fatigue levels. Lastly, the integration of real-time monitoring and feedback mechanisms could be explored to develop adaptive models that respond dynamically to changing conditions. Overall, these future directions aim to refine and extend the current models for more robust and applicable fatigue prediction in practical scenarios.

References

Ansari, S., Du, H., Naghdy, F. and Stirling, D. (2022). Automatic driver cognitive fatigue detection based on upper body posture variations, *Expert Systems with Applications* **203**: 117568–117568.

Ansari, S., Naghdy, F., Du, H. and Pahnwar, Y. N. (2021). Driver mental fatigue detection based on head posture using new modified reLU-BiLSTM deep neural network, *IEEE Transactions on Intelligent Transportation Systems* **23**(8): 10957–10969.

Chinoy, E. D., Cuellar, J. A., Jameson, J. T. and Markwald, R. R. (2022). Performance of four commercial wearable sleep-tracking devices tested under unrestricted conditions at home in healthy young adults, *Nature and Science of Sleep* pp. 493–516.

Choi, B., Jebelli, H. and Lee, S. (2019). Feasibility analysis of electrodermal activity (EDA) acquired from wearable sensors to assess construction workers' perceived risk, *Safety science* **115**: 110–120.

Cui, Y., Zhang, M., Li, J., Luo, H., Zhang, X. and Fu, Z. (2019). WSMS: Wearable stress monitoring system based on IoT multi-sensor platform for living sheep transportation, *Electronics* **8**(4): 441–441.

Guk, K., Han, G., Lim, J., Jeong, K., Kang, T., Lim, E. K. and Jung, J. (2019). *Evolution of Wearable Devices with Real-Time Disease Monitoring for Personalized Healthcare. Nanomaterials (Basel)* **9**: 813–813.

Hooda, R., Joshi, V. and Shah, M. (2021). A comprehensive review of approaches to detect fatigue using machine learning techniques, *Chronic Diseases and Translational Medicine* .

Li, Y., Wang, D. and Liu, F. (2022). The auto-correlation function aided sparse support matrix machine for eeg-based fatigue detection, *IEEE Transactions on Circuits and Systems II: Express Briefs* **70**(2): 836–840.

Liu, P., Chi, H. L., Li, X. and Guo, J. (2021). Effects of dataset characteristics on the performance of fatigue detection for crane operators using hybrid deep neural networks, *Automation in Construction* **132**: 103901–103901.

Ma, L., Ma, H., Zhan, X. and Wang, Y. (2023). *How Do Problem-Solving Demands Influence Employees' Thriving at Work: An Explanation Based on Cognitive Appraisal* .

Marandi, R. Z., Madeleine, P., Omland, Ø., Vuillerme, N. and Samani, A. (2019). An oculometrics-based biofeedback system to impede fatigue development during computer work: A proof-of-concept study, *PLoS One* **14**(5): 213704–213704.

Oliver, L. S., Sullivan, J. P., Russell, S., Peake, J. M., Nicholson, M., Mcnulty, C. and Kelly, V. G. (2021). Effects of nutritional interventions on accuracy and reaction time with relevance to mental fatigue in sporting, military, and aerospace populations: a systematic review and meta-analysis, *International Journal of Environmental Research and Public Health* **19**(1): 307–307.

Parekh, V., Shah, D. and Shah, M. (2020). Fatigue detection using artificial intelligence framework, *Augmented Human Research* **5**: 1–17.

Santhosh, S. and Ar, P. K. (2022). Auto mobilist Fatigue Detection Using Machine Learning Algorithms, *2022 1st International Conference on Computational Science and Technology (ICCST)* pp. 467–471.

Shin, J., Kim, S., Yoon, T., Joo, C. and Jung, H. I. (2019). Smart Fatigue Phone: Real-time estimation of driver fatigue using smartphone-based cortisol detection, *Biosensors and Bioelectronics* **136**: 106–111.

Souchet, A. D., Lourdeaux, D., Pagani, A. and Rebenitsch, L. (2023). A narrative review of immersive virtual reality's ergonomics and risks at the workplace: cybersickness, visual fatigue, muscular fatigue, acute stress, and mental overload, *Virtual Reality* **27**(1): 19–50.

Tuncer, T., Dogan, S., Ertam, F. and Subasi, A. (2021). A dynamic center and multi threshold point based stable feature extraction network for driver fatigue detection utilizing EEG signals, *Cognitive neurodynamics* **15**: 223–237.