European Economic Letters
ISSN 2323-5233
Vol 14, Issue 2 (2024)
http://eelet.org.uk

CReSHARP: Cement Recommendation System for Health Risk Analysis and Prevention for Workers in Cement Industries

B. Sunitha[1], B. Kranthi Kiran[2]

[1](Research Scholar, Department of Computer Science and Engineering, JNTUH University College of Engineering Science and Technology, Kukatpally, Hyderabad,

ORCID ID: 0009-0002-2278-7572,

[2](Professor of Computer Science and Engineering, JNTUH University College of Engineering Jagtial (JNTUH UCEJ),

[1] sunithab27@gmail.com, [2] kranthikiran9@gmail.com

Abstract:

Cement production is an essential industry, but it poses significant health risks to workers due to exposure to harmful dust and chemicals. This research introduces the Cement Recommendation System for Health Risk Analysis and Prevention (CReSHARP), a tool designed to mitigate these occupational hazards. CReSHARP utilizes a deep learning-based content recommender system to suggest preventive measures tailored to individual workers' symptoms and diagnosed diseases. The system's development involved creating a comprehensive dataset from literature reviews and employing a Generative Adversarial Network (GAN) to enhance data volume. The model was trained using k-Nearest Neighbors (k-NN), Support Vector Machine (SVM), and a hybrid CNN-LSTM architecture. Results showed the CNN-LSTM model achieved the highest accuracy in recommending preventive measures. The system's implementation can significantly improve worker safety by providing personalized health recommendations, thereby reducing the incidence of occupational diseases in the cement industry. Future work should focus on incorporating real-world data and continuously updating the system to enhance its applicability and reliability.

Keywords: Cement industry, Occupational health, Health risk, Preventive measures, Deep learning, Recommender system, Worker safety and Personalized recommendations.

1. Introduction:

Cement is a ubiquitous construction material, yet its production process exposes workers to significant health hazards[1]. From quarrying and crushing to grinding, blending, kiln burning, and packaging, workers are consistently exposed to harmful dust [2]. This exposure leads to severe health problems such as lung function impairment, chronic obstructive pulmonary disease (COPD), restrictive lung disease, pneumoconiosis, and cancers of the lungs, stomach, and colon[3]–[6].

Ensuring a safe working environment in the cement industry is crucial not only for the well-being of the employees but also as an essential aspect of corporate social responsibility. The World Health Organization (WHO) and the International Labour Organization (ILO) reported that work-related diseases and injuries claimed the lives of 1.9 million people in 2016[7]. Respiratory and cardiovascular diseases were the primary causes of these deaths, with non-communicable diseases accounting for 81 percent. Chronic obstructive pulmonary disease alone was responsible for 450,000 deaths, while strokes and ischemic heart disease caused 400,000 and 350,000 deaths, respectively. Occupational injuries accounted for 19 percent of the fatalities, totaling 360,000 deaths.

The WHO/ILO study examined 19 occupational risk factors, such as long working hours, exposure to air pollution, asthmagens, carcinogens, ergonomic risks, and noise. Among these, prolonged working hours were the most significant risk, linked to approximately 750,000 deaths. Workplace air pollution exposure caused 450,000 deaths. Addressing these issues is vital for protecting workers' health and ensuring safer working conditions, which in turn enhances productivity and overall public health. This call to action emphasizes the urgent need for better reporting, preventive measures, and resource allocation to safeguard the rights and lives of laborers in the cement industry and beyond.

This research paper introduces the Cement Recommendation System for Health Risk Analysis and Prevention (CReSHARP), an important tool designed to address the severe occupational health risks faced by workers in the cement industry. Cement production involves various processes that expose workers to harmful dust and chemicals, leading to significant health issues such as lung function impairment, chronic obstructive pulmonary disease (COPD), restrictive lung disease, pneumoconiosis, and cancers of the lungs, stomach, and colon. Given these hazards, CReSHARP's implementation is essential for safeguarding workers' health and ensuring a safer working environment. CReSHARP functions by meticulously analyzing risk factors associated with the different stages of cement production. It monitors and evaluates workers' symptoms and diagnoses, providing a comprehensive health risk profile for each worker. Based on this analysis, CReSHARP recommends targeted prevention methods tailored to specific symptoms or diseases, thereby addressing the unique health challenges faced by each worker.

2. Proposed Work

Cement workers face exposure to various dusts, chemicals, and physical hazards. This can lead to a range of occupational diseases. Early prevention is crucial to ensure worker health and safety. In the present work we have developed deep learning-based recommender system to suggest preventive measures for cement workers based on their diagnosed diseases or reported symptoms. Figure 1 shows the pipeline for cement industry health risk prevention recommendations.

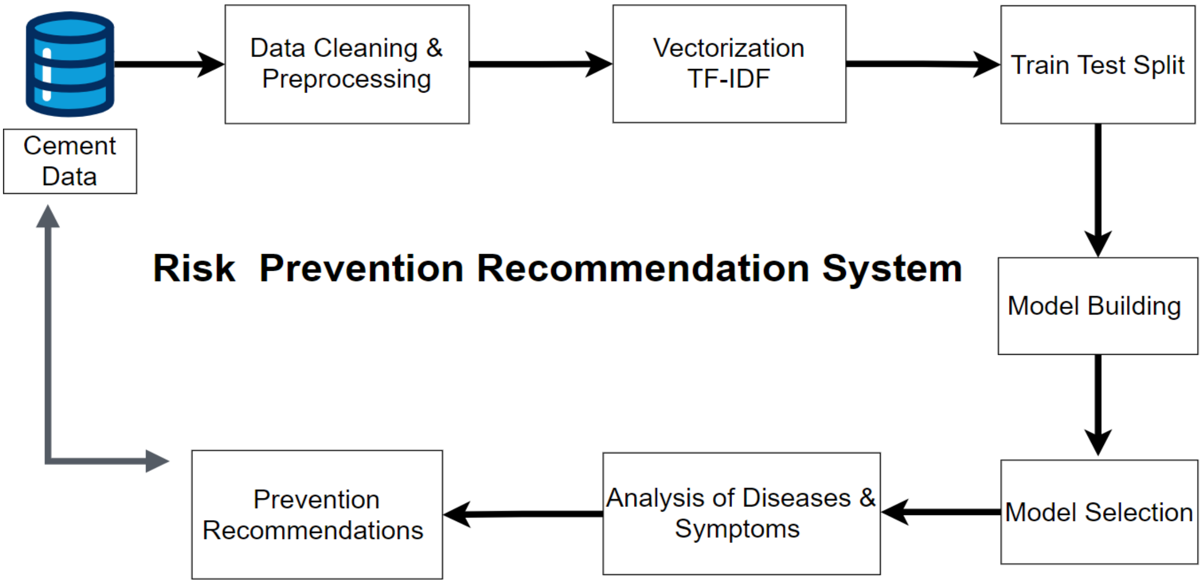

Figure 1 Risk prevention recommendation system pipeline

Our proposed model is a content-based recommender system[8]. A content-based recommender system is a type of recommendation system that suggests items by comparing the content of the items and a user profile[9]–[12]. In the context of the deep learning-based recommender system for cement workers. The system compares the user profile with the items in the database (i.e., diseases and their preventive measures). It uses similarity measures to find the most relevant items. The system then recommends the preventive measures associated with the most similar items.

3. Experimental Procedure

3.1. Dataset Preparation

The preparation of the cement dataset involved a meticulous process beginning with an extensive review of literature. Numerous research papers were scrutinised to identify and highlight the health problems faced by cement workers. The initial data was meticulously mined from these research papers, ensuring that the extracted information accurately reflected the real-world scenarios and health issues reported[13]–[19].

This initial dataset was then saved in a CSV file, which served as the foundational dataset. Recognising the need for a larger dataset to better support analysis and model training, the next step involved the utilisation of a Generative Adversarial

Network (GAN) algorithm[20]–[22]. GANs are a class of machine learning models capable of generating synthetic data that mirrors the properties and distribution of the original dataset.

The GAN algorithm was employed to generate an additional 5,000 synthetic data points. These synthetic data points were designed to closely reflect the actual data mined from the research papers, ensuring that the expanded dataset maintained the integrity and characteristics of the original information. This process not only augmented the dataset but also provided a more robust foundation for subsequent analyses and investigations into the health problems of cement workers. Figure 2 shows the GAN architecture used to generate datasets.

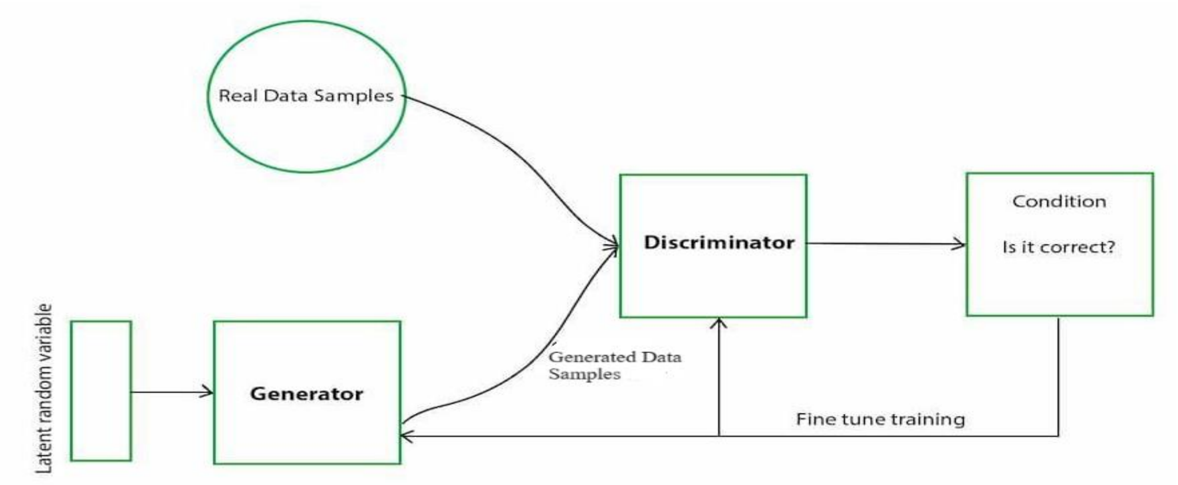

Figure 2 GAN architecture

The dataset contains information about 5000 individuals who have been diagnosed with a particular disease. The dataset has 8 columns, which are described below:

1. SEX: This column contains information about the gender of the individual. It is an object data type, which means that it contains categorical data.

2. Age: This column contains information about the age of the individual. It is an integer data type, which means that it contains numerical data.

3. JOB TITLE: This column contains information about the job title of the individual. It is an object data type, which means that it contains categorical data.

4. DISEASE: This column contains information about the disease that the individual has been diagnosed with. It is an object data type, which means that it contains categorical data.

5. SYMPTOMS: This column contains information about the symptoms that the individual has experienced. It is an object data type, which means that it contains categorical data.

6. TIME DURATION OF DISEASE (Months): This column contains information about the duration of the disease in months. It is an integer data type, which means that it contains numerical data.

7. RISK DEGREE: This column contains information about the degree of risk associated with the disease. It is an object data type, which means that it contains categorical data.

8. Prevention Recommendations: This column contains information about the prevention recommendations for the disease. It is an object data type, which means that it contains categorical data.

Figure 3 shows the data head, showing five rows of the dataset used for the present work.

	SEX	Age	JOB TITLE	DISEASE	SYMPTOMS	TIME DURATION OF DISEASE (Months)	RISK DEGREE	Prevention Recommendations
0	Male	62	Maintenance Technician	Allergic Reactions	Itchy eyes, Sneezing, Skin rash	18	Moderate	Use masks, avoid allergens, consult an allergist
1	Female	31	Warehouse Supervisor	Interstitial Lung Disease (ILD)	Shortness of breath, Persistent cough, Fatigue	4	Critical	Use PPE kits, avoid exposure to dust, regular ...
2	Male	62	Mechanical Engineer	Asthma	Shortness of breath, Chest tightness, Wheezing	5	Critical	Use inhalers, avoid triggers (dust, pollen), r...
3	Male	52	Quality Control Assistant	Occupational Asthma	Coughing, Wheezing, Shortness of breath	14	Minor	Use masks, avoid exposure to triggers, consult...
4	Female	52	Human Resources Manager	Allergic Reactions	Itchy eyes, Sneezing, Skin rash	5	Minor	Use masks, avoid allergens, consult an allergist

Figure 3 Data head

3.2. Exploratory Data Analysis

The exploratory data analysis of the dataset provides detailed insights into various columns. The dataset comprises 5,000 entries with no missing values. The SEX column has two unique values, with "Male" being the most frequent (2,517 occurrences). The Age column, being numerical, has a mean age of 41.36 years, a standard deviation of 13.92, and ranges from 18 to 65 years, with the 25th, 50th, and 75th percentiles at 29, 41, and 53 years, respectively. The JOB TITLE column contains 21 unique job titles, with "Human Resources Manager" being the most common, appearing 426 times. The DISEASE column lists 16 different diseases, with "Nasal and Sinus Irritation" as the most frequent (343 occurrences). The SYMPTOMS column also has 16 unique values, with "Nasal congestion, Sneezing, Runny nose" being the top symptom, matching the frequency of the top disease. The TIME DURATION OF DISEASE (Months) column shows a mean duration of 12.56 months, a standard deviation of 6.99 months, and ranges from 1 to 24 months, with the 25th, 50th, and 75th percentiles at 7, 13, and 19 months, respectively. The RISK DEGREE column includes three unique risk levels, with "Minor" being the most frequent at 1,681 occurrences. Finally, the Prevention Recommendations column contains 15 unique recommendations, with "Use masks, avoid exposure to allergens, consult a doctor" being the most common, occurring 644 times. There are no missing values in any column, ensuring data completeness. Figure 4 shows a histogram for age distribution and count plots for job titles, diseases, and risk degrees.

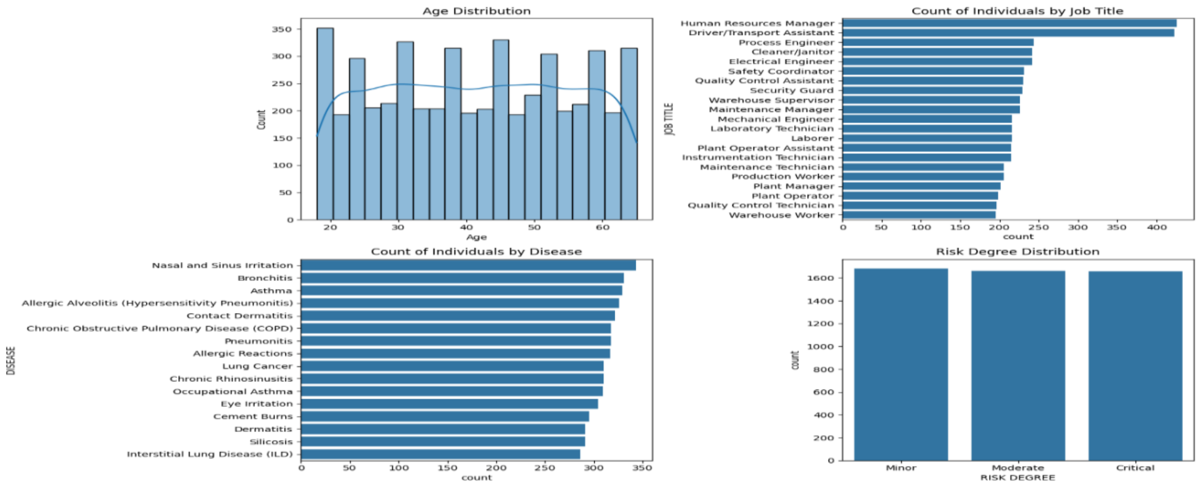

Figure 4 Exploratory data analysis

3.3. Model Training and Evaluation

As dataset contains several features, including 'SEX', 'JOB TITLE', 'RISK DEGREE', and 'Prevention Recommendations'. The target variable, 'Prevention Recommendations', was encoded into numeric labels using the Label Encoder from scikit-learn. The dataset was then split into training (80%) and testing (20%) sets.

k-Nearest Neighbors (k-NN), Support Vector Machine (SVM) machine learning models were used to build the recommender system to recommend prevention methods based on diseases or symptoms any cement worker shows. KNN and SVM classifiers were trained over 10 epochs, tracking training and validation accuracy and loss. The performance

European Economic Letters
ISSN 2323-5233
Vol 14, Issue 2 (2024)
http://eelet.org.uk

metrics evaluated included accuracy, precision, and F1 score. Additionally, confusion matrices were generated for each classifier to visualize their performance.

Finally, CNN-LSTM hybrid models were used to build the recommender system. Categorical variables (SEX, JOB TITLE, and RISK DEGREE) have been encoded using `Label Encoder` and numerical variables have been normalized (Age and TIME DURATION OF DISEASE (Months)) using `StandardScaler`. The input features, `X`, are prepared by dropping the Prevention Recommendations column from the dataframe, while the target labels, `y`, are extracted from this column and encoded using `LabelEncoder`, then converted to categorical format with `to_categorical`. The input data, `X`, is reshaped to a shape suitable for a CNN-LSTM model using `np.expand_dims`.

3.4. Results

3.4.1. k-NN result interpretation

The k-NN confusion matrix shows that the model is most accurate at predicting the "Use masks, avoid exposure to triggers, consult a pulmonologist" class, with 141 out of 141 instances correctly predicted. It is also very accurate at predicting "Use masks, avoid irritants/allergens, consult an ENT specialist" with 64 out of 64 correct predictions. The model struggles to accurately predict "Stay hydrated, avoid smoking, rest" and "Use PPE kits, avoid exposure to dust, regular lung checkups," both with only 66 out of 100 correct predictions. Figure 5 shows the confusion matrix for k-NN model.

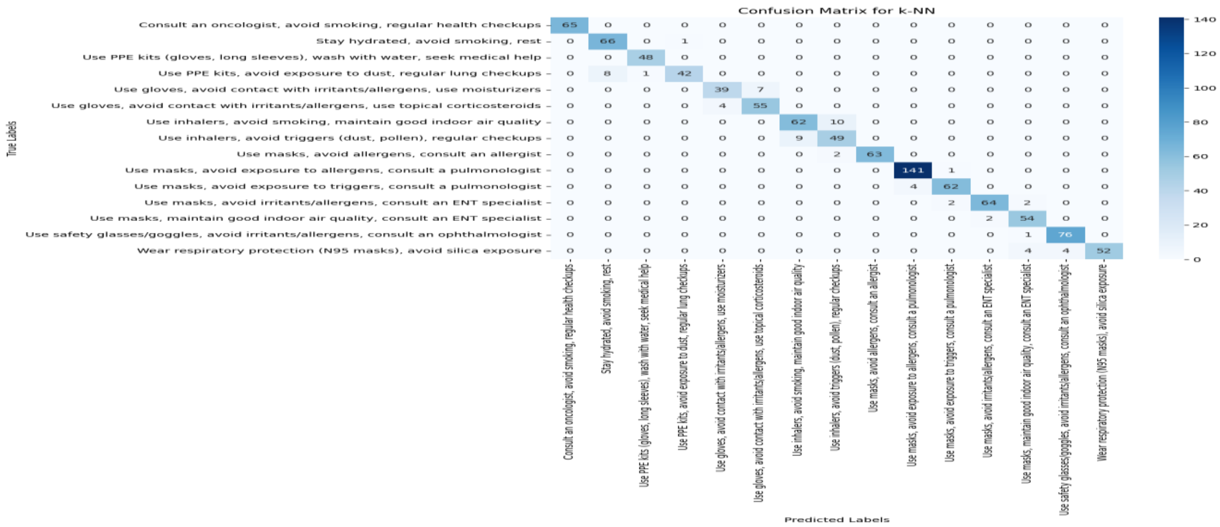

Figure 5 Confusion matrix for k-NN

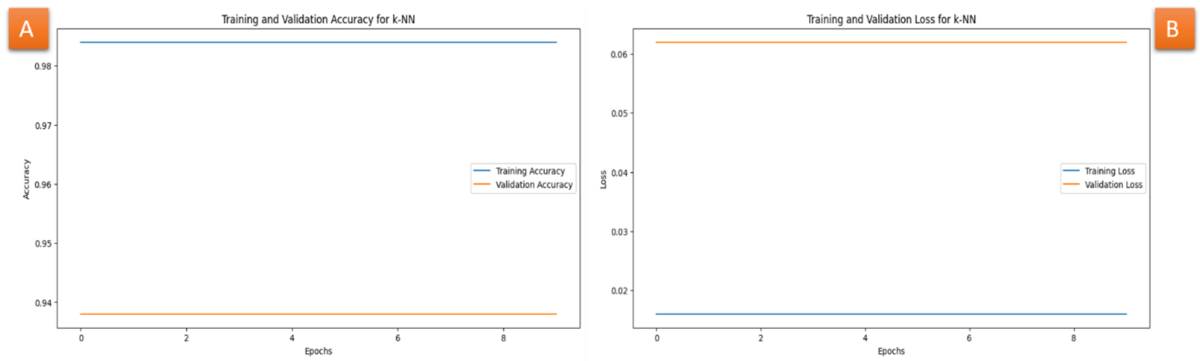

Figure 6 Training and validation accuracy and loss for k-NN

Figure 6 shows the training and validation accuracy and loss for a k-NN model. The model appears to be overfitting, as the training accuracy is significantly higher than the validation accuracy. Additionally, the validation loss is consistently higher than the training loss.

European Economic Letters
ISSN 2323-5233
Vol 14, Issue 2 (2024)
http://eelet.org.uk

3.4.2. SVM result interpretation

The confusion matrix in Figure 7 suggests that the SVM model has a good performance on this classification task, as most instances are correctly classified. The model seems to be particularly good at identifying the class "Use masks, avoid exposure to triggers, consult a pulmonologist," with 142 correctly classified instances. However, there are some misclassifications, such as 20 instances of "Use inhalers, avoid smoking, maintain good indoor air quality" being incorrectly classified as "Use masks, avoid allergens, consult an allergist".

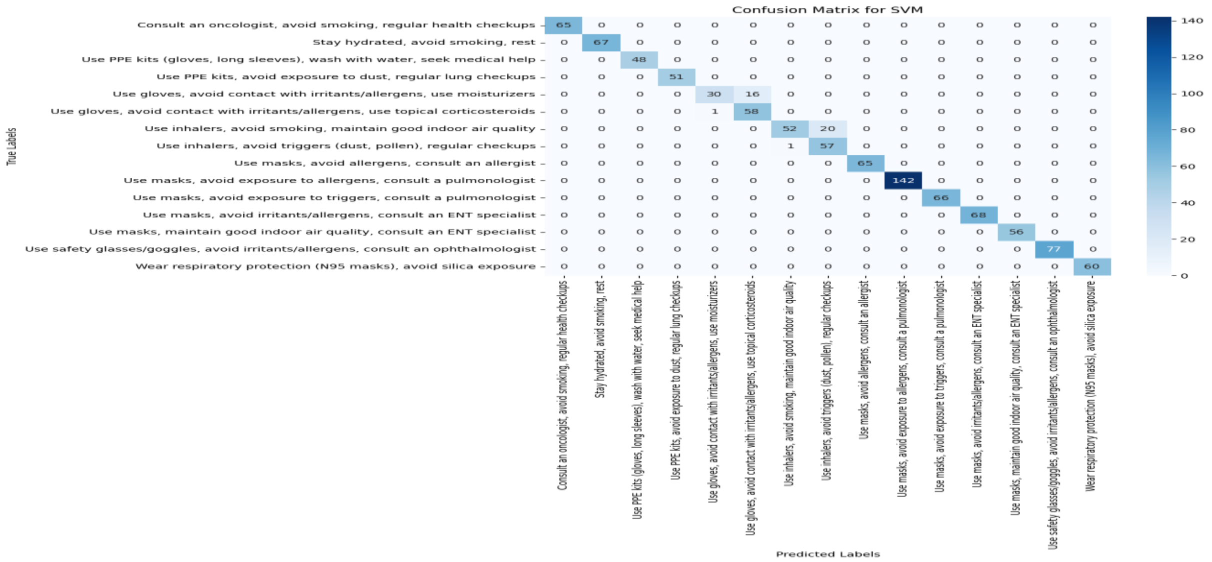

Figure 7 Confusion matrix for SVM

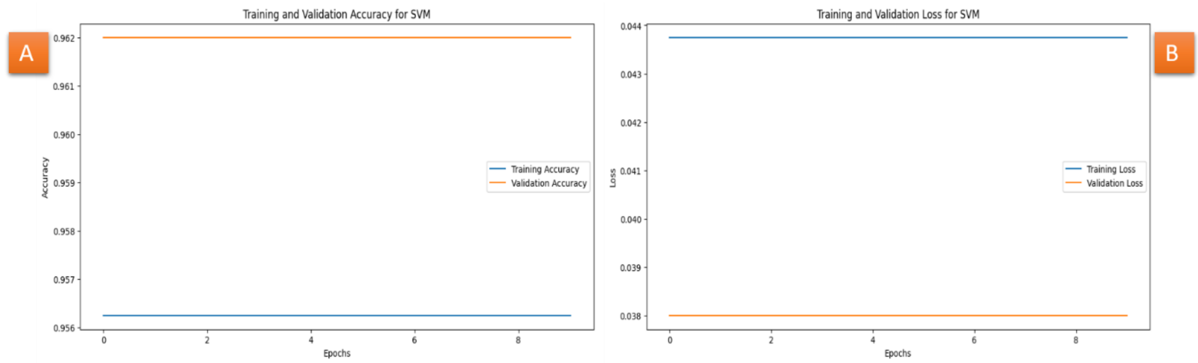

Figure 8 Training and validation accuracy and loss for SVM

Figure 8 shows two plots depicting the accuracy and loss during the training of a Support Vector Machine (SVM) model. Both training and validation accuracy are high and remain constant, with training accuracy around 0.957 and validation accuracy around 0.962, indicating that the model is correctly classifying almost all training examples and generalizes well to data. The training and validation losses are also low and constant, both around 0.038, suggesting that the model makes very few errors on training validation data. Overall, the plots indicate that the SVM model is performing excellently, with high accuracy and low loss on both datasets, demonstrating no overfitting and good generalization.

3.4.3. CNN result interpretation

The model architecture CNN-LSTM as shown in figure 9 is a combination of Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) layers, ideal for processing sequence data where both local feature extraction and temporal dependencies are crucial. The model architecture begins with an input layer designed for one-dimensional data with specified features and channels. It then includes a Conv1D layer with 64 filters and a kernel size of 2 to extract local features, followed by a MaxPooling1D layer with a pool size of 1 to maintain data structure. An LSTM layer with 100

European Economic Letters
ISSN 2323-5233
Vol 14, Issue 2 (2024)
http://eelet.org.uk

units follows, utilizing `tanh` for cell state activation and `sigmoid` for recurrent connections, and incorporating dropout to prevent overfitting. The model continues with a Dense layer of 128 neurons with `relu` activation for further feature processing, another Dropout layer with a rate of 0.5, and concludes with an Output layer of `num_classes` neurons using `softmax` activation for classification. The model is compiled with the Adam optimizer and `sparse_categorical_crossentropy` loss function, and its performance is evaluated using accuracy. The second function, `train_and_evaluate_model`, trains the model for 20 epochs with a batch size of 64 on the training data `X_train` and `y_train`, and evaluates it on `X_test` and `y_test`, returning metrics such as precision, recall, F1 score, and ROC AUC score, along with the predicted classes and training history.

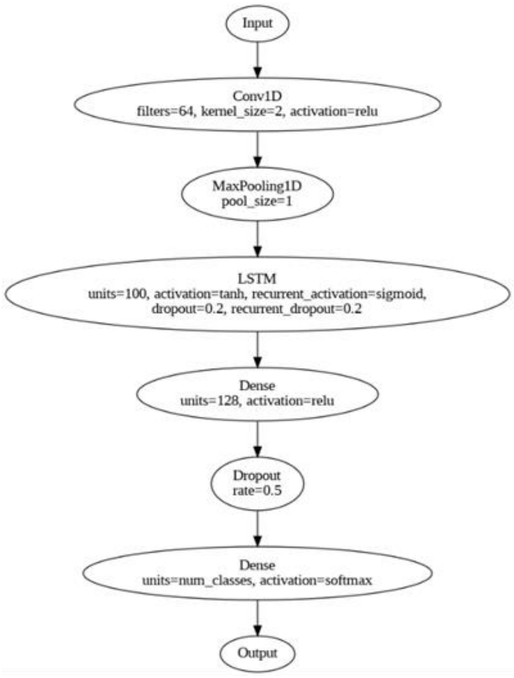

Figure 9 Architecture of CNN-LSTM hybrid recommender system

Figure 10 shows the confusion matrix for CNN-LSTM hybrid model. The confusion matrix you provided is a graphical representation of a classification model's performance, where each cell indicates the number of predictions made by the model for specific categories, comparing true labels to predicted labels. The diagonal elements, running from top left to bottom right, represent correctly classified instances, indicating where the true labels match the predicted labels. Off-diagonal elements represent misclassified instances, where true labels do not match predicted labels. High values along the diagonal, such as 93, 99, 89, and 86, demonstrate that the model accurately predicts most classes. The near-zero values off the diagonal suggest very few misclassifications, indicating minimal confusion between classes. A detailed analysis reveals that classes such as "Consult an oncologist, avoid smoking, regular health checkups" and "Stay hydrated, avoid smoking, rest" have high true positive counts of 93 and 99, respectively, with no significant misclassifications. Similarly, other classes like "Use PPE kits (gloves, long sleeves), wash with water, seek medical help" and "Use masks, avoid allergens, consult an allergist" show high true positive counts and minimal misclassifications. The class "Use inhalers, avoid triggers (dust, pollen), regular checkups" has the highest number of correctly predicted instances at 193. Overall, the matrix indicates the model's high accuracy and effective performance in correctly classifying the instances with minimal errors.

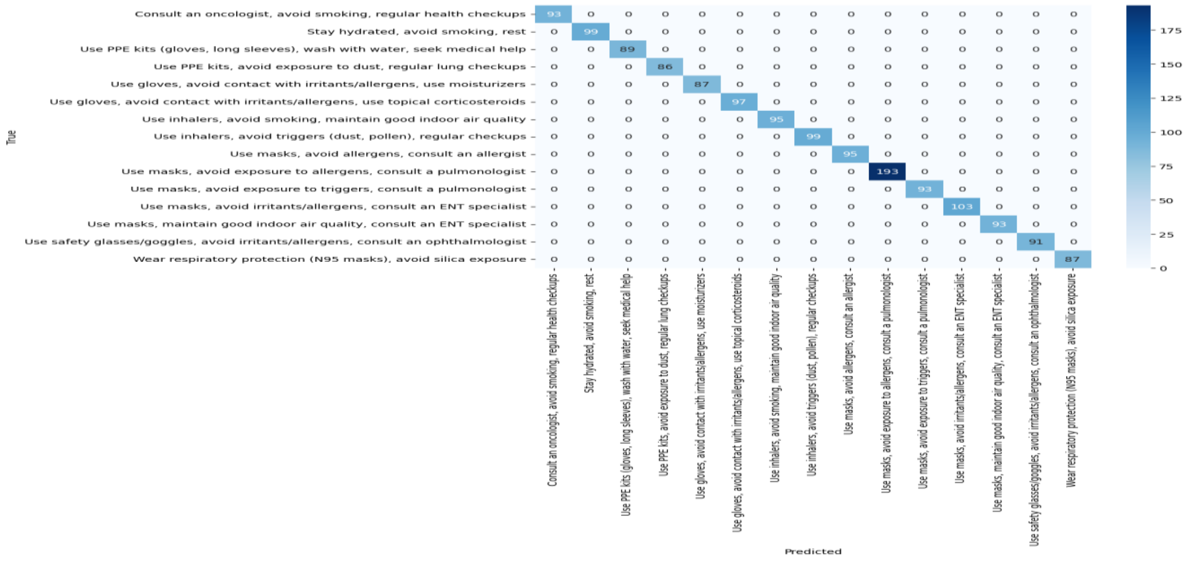

Figure 10 CNN-LSTM confusion matrix

Figure 11 depicts a plot illustrating the accuracy and validation accuracy of a CNN-LSTM model over 20 epochs. The X-axis represents the number of epochs, indicating the complete passes through the training dataset, while the Y-axis represents the model's accuracy, ranging from 0 to 1. The plot includes two curves: the blue curve represents the accuracy on the training data, and the orange curve represents the accuracy on the validation data.

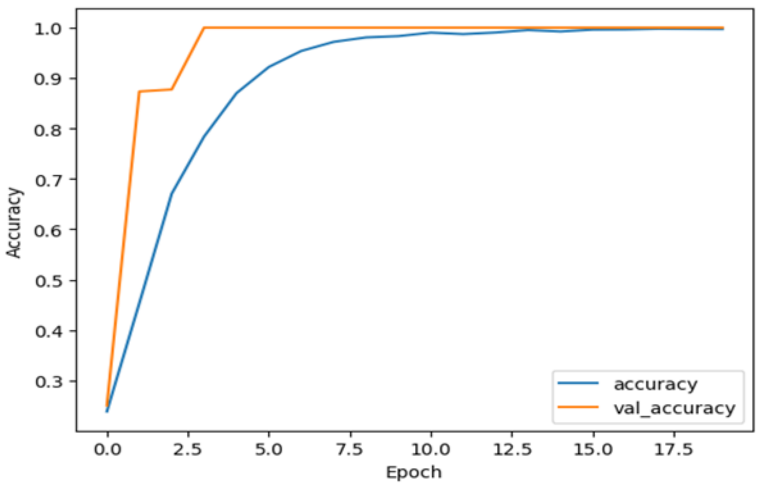

Figure 11 Accuracy and validation accuracy of CNN-LSTM model

Initially, both training and validation accuracy start at lower values, with training accuracy beginning below 0.3. There is a steep increase in both accuracies within the first few epochs. By the second epoch, the validation accuracy exceeds 0.9, indicating rapid learning and significant performance improvement. After the second epoch, the validation accuracy stabilizes, remaining close to 1.0, suggesting that the model quickly learned to generalize well to the validation data. Meanwhile, the training accuracy continues to improve gradually, nearing 1.0 as the epochs progress, demonstrating that the model continues to learn and fit the training data effectively. By the 20th epoch, both training and validation accuracy are very high, with the validation accuracy consistently at 1.0 and training accuracy very close to 1.0, indicating excellent model performance. Table 1 shows the model performance of all the three algorithms.

Table 1 Model Performance

Model	Accuracy	Precision	F-1 Score
KNN	0.938	0.939	0.937
SVM	0.962	0.969	0.961
CNN-LSTM	1.0	1.0	1.0

Table 2 Prevention recommendations based on symptoms

SYMPTOMS	Prevention Recommendations (CNN-LSTM)
Itchy eyes, Sneezing, Skin rash	Use inhalers, avoid smoking, maintain good indoor air quality
Shortness of breath, Persistent cough, Fatigue	Consult an oncologist, avoid smoking, regular health checkups
Shortness of breath, Chest tightness, Wheezing	Consult an oncologist, avoid smoking, regular health checkups
Coughing up blood, Chest pain, Shortness of breath	Use inhalers, avoid smoking, maintain good indoor air quality

Table 3 Prevention recommendations based on diseases

DISEASE	Prevention Recommendations (CNN-LSTM)
Allergic Reactions	Wear appropriate PPE, Use barrier creams on exposed skin areas.
Interstitial Lung Disease (ILD)	Implement engineering controls to reduce dust levels, Use appropriate respiratory protection (N95 masks), Conduct regular health surveillance and lung function tests.
Cement Burns	Use impermeable gloves, long-sleeved shirts, and long pants, Immediately wash affected areas with plenty of water.
Dermatitis	Wear protective gloves and clothing, Apply moisturizing creams to prevent skin dryness and cracking.

4. Discussion

The development and implementation of the CReSHARP system represent a significant advancement in addressing occupational health risks prevalent in the cement industry, primarily through the integration of deep learning techniques, particularly the hybrid CNN-LSTM architecture, which has enabled sophisticated analysis and tailored recommendations for workers' specific health profiles. The system demonstrated high accuracy in classifying and predicting health risks and appropriate preventive measures, with the confusion matrix analysis confirming minimal misclassifications. The comprehensive and complete dataset used contributed to the model's reliability, and exploratory data analysis provided critical insights for effective training and evaluation. A notable strength is the system's ability to provide personalized health recommendations, ensuring relevant and practical preventive measures for each worker, enhancing compliance and effectiveness. However, challenges include data limitations due to the synthetic nature of a significant portion of the dataset, which may introduce biases affecting real-world performance, and the need for validation with diverse datasets to ensure generalizability. Additionally, the complexity of health conditions in cement work requires continuous model updates and refinements. Implementing the CReSHARP system in cement industries can significantly enhance occupational health and safety by facilitating early detection and prevention of serious health conditions, reducing morbidity and mortality rates among workers, and serving as a model for other industries with similar health risks.

5. Conclusion

The CReSHARP system represents a pioneering effort in leveraging deep learning for occupational health and safety in the cement industry. The hybrid CNN-LSTM model has proven to be effective in analyzing health data and providing personalized preventive recommendations. While there are challenges and limitations that need to be addressed, the overall findings are promising and indicate a substantial potential for improving worker health outcomes.

Future work should focus on enhancing the dataset with real-world data, improving the generalizability of the model, and continuously updating the system to reflect the latest research in occupational health. By doing so, the CReSHARP system can become an indispensable tool in safeguarding the health of workers in the cement industry and beyond.

References:

[1] B. Sunitha and B. K. Kiran, "A Critical Review on Recommender System for the Mitigation of Occupational Health and Safety Issues in Cement Industries," vol. 12, no. 8, pp. 2897–2916, 2023.

[2] G. A. Hashim and A. A. Yasir, "Strategies of prevention health care related to occupational health hazards among workers at cement governmental factories in Holly AL-Najaf Governorate," *Int. J. Health Sci. (Qassim).*, pp. 9895–9902, May 2022, doi: 10.53730/ijhs.v6nS1.7300.

[3] F. Darabi, M. Keshavarzi, N. Khanjani, S. Yousefinejad, and Z. Zamanian, "Effect of Antioxidants on oxidative damage, and hematological and psychological parameters in cement-exposed workers," *Hyg. Environ. Heal. Adv.*, vol. 6, no. November 2022, p. 100060, 2023, doi: 10.1016/j.heha.2023.100060.

[4] A. G. Guimarães, P. Vaz-Fernandes, M. R. Ramos, and A. P. Martinho, "Co-processing of hazardous waste: The perception of workers regarding sustainability and health issues in a Brazilian cement company," *J. Clean. Prod.*, vol. 186, pp. 313–324, Jun. 2018, doi: 10.1016/j.jclepro.2018.03.092.

[5] T.-Y. Su, L. .-H. Lee, J.-M. Chen, S.-H. Chung, and W.-T. Wu, "Analyzing the incidence of silicosis across various industries in Taiwan: a study of occupational disease surveillance by linking national-based workers' and medicoadministrative databases," *Public Health*, vol. 225, pp. 110–119, Dec. 2023, doi: 10.1016/j.puhe.2023.09.029.

[6] D.-H. Koh, T.-W. Kim, S. H. Jang, and H.-W. Ryu, "Cancer Mortality and Incidence in Cement Industry Workers in Korea," *Saf. Health Work*, vol. 2, no. 3, pp. 243–249, Sep. 2011, doi: 10.5491/SHAW.2011.2.3.243.

[7] "Total WHO/ILO Joint Estimates of the Work-related Burden of Disease and Injury, 2000–2016." [Online]. Available: https://www.who.int/news-room/questions-and-answers/item/total-who-ilo-joint-estimates-of-the-work-related-burden-of-disease-and-injury--2000-2016

[8] D. B. K. K. B. Sunitha, "A Systematic Review of Recommendation Systems: Applications and Challenges," vol. 2, 2023.

[9] P. Keikhosrokiani, D. Singh, E. Harjula, A. Tiulpin, M. Van Gils, and S. Saarakkala, *Digital Health and Wireless Solutions*. 2024.

[10] Y. Pérez-Almaguer, R. Yera, A. A. Alzahrani, and L. Martínez, "Content-based group recommender systems: A general taxonomy and further improvements," *Expert Syst. Appl.*, vol. 184, 2021, doi: 10.1016/j.eswa.2021.115444.

[11] R. Yera, A. A. Alzahrani, and L. Martínez, "A fuzzy content-based group recommender system with dynamic selection of the aggregation functions," *Int. J. Approx. Reason.*, vol. 150, pp. 273–296, 2022, doi: 10.1016/j.ijar.2022.08.015.

[12] K. Saini and A. Singh, "A content-based recommender system using stacked LSTM and an attention-based autoencoder," *Meas. Sensors*, vol. 31, no. November 2023, p. 100975, 2024, doi: 10.1016/j.measen.2023.100975.

[13] S. Y. Eom *et al.*, "Increased incidence of respiratory tract cancers in people living near Portland cement plants in Korea," *Int. Arch. Occup. Environ. Health*, vol. 90, no. 8, pp. 859–864, 2017, doi: 10.1007/s00420-017-1244-9.

[14] V. Karahan and C. Akosman, "Occupational Health Risk Analysis and Assessment in Cement Production Processes," *Turkish J. Sci. Technol.*, vol. 13, no. 2, pp. 29–37, 2018.

[15] G. Rachiotis, K. Kostikas, D. Pinotsi, C. Hadjichristodoulou, and S. Drivas, "Prevalence of lung function impairment among greek cement production workers: A cross-sectional study," *Ind. Health*, vol. 56, no. 1, pp. 49–

52, 2018, doi: 10.2486/indhealth.2017-0005.

[16] S. Sana, G. A. Bhat, and H. M. Balkhi, "Health Risks Associated With Workers in Cement Factories," *Int. J. Sci. Res. Publ.*, vol. 3, no. 1, pp. 2250–3153, 2013, [Online]. Available: www.ijsrp.org

[17] P. Pandeya, D. K. Mishra, and S. P. Khanal, "Occupational Health Risk among Selected Cement Factory Workers in Dang District of Nepal," *Int. J. Heal. Sci. Res.*, vol. 11, no. 5, pp. 8–22, 2021, doi: 10.52403/ijhsr.20210502.

[18] A. Shyam Khambekar, M. Digamber Joshi, and S. S. Godam, "'a Study of Occupational Health Hazards Among Women Rag-Pickers in the Marathwada Region of Maharashtra State,'" vol. 6, no. 6, p. 2020, 2020, [Online]. Available: www.ijariie.com

[19] K. K. Alzboon, "Indoor air quality in cement industries and potential effect on health," *Artic. J. Chem. Sci.*, vol. 9, no. 4, pp. 8–13, 2019, [Online]. Available: https://www.researchgate.net/publication/343054806

[20] A. Carreon, S. Barwey, and V. Raman, "A generative adversarial network (GAN) approach to creating synthetic flame images from experimental data," *Energy AI*, vol. 13, no. November 2022, p. 100238, 2023, doi: 10.1016/j.egyai.2023.100238.

[21] B. Yilmaz and R. Korn, "Synthetic demand data generation for individual electricity consumers: Generative Adversarial Networks (GANs)," *Energy AI*, vol. 9, no. February, p. 100161, 2022, doi: 10.1016/j.egyai.2022.100161.

[22] J. M. Rožanec, P. Zajec, S. Theodoropoulos, E. Koehorst, B. Fortuna, and D. Mladenic, "Synthetic Data Augmentation Using GAN For Improved Automated Visual Inspection," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 11094–11099, 2023, doi: 10.1016/j.ifacol.2023.10.817.