REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No. 1 (2024)
http://www.veterinaria.org
Article Received: Revised: Accepted:

Modelling Effective Parameters For Enhanced Interaction Practices In Industry: An AI Perspective

Ashwini Kumar[1*], Rekha Agarwal[2], Archana Singh[3]

[1*]Amity Institute of Information Technology, Amity University, Noida, Uttar Pradesh, 201301, India, mrashwinikumar@gmail.com
[2]AIIT, Amity University, Noida, Uttar Pradesh, 201301, India, ragarwal@amity.edu
[3]Caliper, Foresight Health Solutions LLC, Noida, Uttar Pradesh, 201301, India, archana.elina@gmail.com

***Corresponding Author:** Ashwini Kumar
*Amity Institute of Information Technology, Amity University, Noida, Uttar Pradesh, 201301, India, mrashwinikumar@gmail.com

ABSTRACT
The focal point in digitalization and automation opportunities has been improving the industry's interaction practices for better productivity, customer engagement, and operational efficiency. This study, titled "Modelling Effective Parameters for Enhanced Interaction Practices in Industry: The paper titled "An AI Perspective", discusses how Interaction parameters by using machine learning methods RF and K-Means on Interaction coups and applying the COFI framework, namely Context, Content, Competency, and Culture. Applying the novel approach of data preprocessing, modelling and evaluation techniques for the utilization of the combination of supervised and unsupervised learning, this study established that the employed supervised models of learning offered an accuracy of 54% as found in the RF model, offering the highest efficient predictive accuracy and generalizable for industrial applications of general understanding. Nevertheless, K-Means clustering shows better results in the case of the COFI metric within the human-AI interaction analysis, regarding the Completeness (CC), the Correctness (CU), and the Accuracy (CA) measures, scoring more than 70%. These results show that even if RF is appropriate for decision-making for fine-tuned predictions, K-Means is better for segmenting targets based on the COFI framework to offer systematic data on interaction. So, this research adds to the limited body of literature on using AI for industrial interaction practices while arguing for integrating both RF and K-Means models to address the diverse needs of today's industries.

Keywords
The COFI framework, the Random Forest algorithm, K Means clustering, interactions, predictive models, Human AI Interactions, industrial mathematics optimization, segmentation analysis, AI typical system models.

Introduction

We must consider the value of effective and dynamic communication practices toward developing automation and digitalization in the industrial era [5]. The novel evolution of technologies such as AI and Machine Learning (ML) has created new opportunities for improving the efficiency and orientation of these interaction practices. Today, businesses rely more on modelling effective interaction within their industries' systems, customers, and general operations, providing the proper context for research in this field. This study, titled Modelling Effective Parameters for Enhanced Interaction Practices in Industry: An AI Perspective, attempts to fulfil these emerging requirements by primarily embracing operational aspects related to enhanced interaction paradigms with enhanced AI approaches, especially regarding existing probabilistic parameters within industrial environments [1]. Over time, industries have embraced big data, encouraging researchers to develop models capable of accommodating the enhanced data information system. In particular, traditional approaches to implementation interaction practices in industries have mainly relied on pre-ordained scripts and structural setups of sectors, which can quickly fail to respond to the dynamic nature of customer needs and the changes in the market environment. This has changed with the transition to machine learning and AI because these new technologies allow for the creation of models that can be trained on data on behaviour and that will, at that, be able to recognize new patterns and then accurately predict future interaction. Such changes are most evident in industries including manufacturing, finance, healthcare and retail since interaction systems, for example, those used in workforce management, can improve productivity and efficiency and, on the consumer end, improve customer experience and satisfaction levels at the heart of adaptive interactions systems A significant component of this publication is the COFI model which is an acronym for Context, Content Competency and Culture. COFI framework was initially developed to offer a holistic view of controlling interaction dynamics or IODs, specifically in organizational or industrial interfaces. Five key factors emerge that define the interaction practices. Each element of the COFI framework plays a distinct role [12]:

- This refers to the situation under which two or more people communicate, specifying how different measures will work in certain situations.

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No. 1 (2024)
http://www.veterinaria.org
Article Received: Revised: Accepted:

- These concern the nature of information that flows in the communication process and affect the frequency and richness of interactions.
- Competency is the skills or ability needed to interact or interconnect, which speaks of expertise or aptitude.
- Culture refers to beliefs, knowledge, and behaviour patterns guiding people within a particular industry.

The benefits of the COFI parameters have been demonstrated in a structured approach to focus on these aspects, which helps comprehend the dynamics and details of interaction in different industries. This work utilizes the COFI framework to assess the improvement of interaction practices through the impact of AI models based on these parameters. Various methodologies of managing interactions are usually keen to evaluate these parameters separately and not holistically, leading to the exploration of systematic pictures that maybe far from the truth in real-life situations. However, this research intends to present a comprehensive view by combining the proposed COFI framework with AI to understand how various interaction parameters interact and what specific business domain requirements should be followed by AI-based models.

This research benefits from using supervised learning models similar to the Random Forest (RF) because of the ability to analyze large datasets and generate accurate predictions. The advantages of using the RF model include higher accuracy and the absence of complexity, which attacks high-dimensional space using an ensemble of decision trees. This makes constructing the RF model ensemble method suitable for reducing biases obtained from the individual trees in areas where the model's accuracy is critical. In this research, the RF model has been used to assess interaction practices. It demonstrates its ability to learn from essential metrics about industrial encounters such as customer inquiries, employee replies, and transaction records. The high accuracy obtained by the proposed RF model in this study demonstrates the model's capability to capture interaction dynamics and the possibility of real-time improvement of engagement practices. The present work also explores the potential of unsupervised learning for classifying the engagement pattern using the K-means clustering technique. K-means clustering is a non-parametric technique that classifies similar data points into different groups without any training data. This method creates a way of categorizing different interaction types and responding to them more desirably and appropriately [13]. For example, in an industrial environment, K-means clustering may cluster customers' queries by topics and thus assist the service department in knowing which issues to address critically. In addition, applying K-means clustering when used with the COFI framework helps determine other areas of convergence in terms of interaction, which enhances the alignment of the overall industrial practices with customer needs and delivered service quality. While using the supervised RF model and the unsupervised K-means clustering method, its comparative analysis is based on the assumption of different aspects of interaction practices. Even though the RF model is accurate and has a high level of precision for predictive tasks, K-means clustering helps discover hidden relationships while deciding the number of clusters essential to represent the structure of interactions. As a result of the balanced integration of these two approaches, the study will offer an integrated picture of interaction practices and evaluate AI models' performance in prediction and clustering indices. This dual approach also fits the goals of the COFI framework since it allows for capturing the quantity and quality of the IPs through the interaction parameters. This study's results would help extend the existing knowledge concerning the role of AI in enhancing interaction practices in industrial settings. The efficiency of the application described in terms of precision, correspondence to the competency's requirements, and customer requirements shows the electiveness of the AI approach for transforming the conventional frameworks to data-driven oriented frameworks tailored to the modern requirements of businesses [9].

Furthermore, comparing the results of both RF and K-means models in the contexts of COFI allows for identifying the potential for further research of supervised and unsupervised approaches of the broad industry. In conclusion, this research emphasizes the importance of defining effective parameters for improving interaction practices in the industry using AI-based methodologies. Allowing for the evaluation of a set of factors influencing the interaction effectiveness as well as the practical guidance regarding the implementation of AI in various industries, the COFI framework integrated with the machine learning model. Thus, we hope that our contribution will make it possible to address the emerging demands and challenges of interactions in the context of the industrial practices of a new digital Data Age [2].

Methodology

The framework adopted for this research paper is an elaborate process for data preprocessing, model selection, comparison, and cross-validation with some constant tuning of the selected model. These steps were chosen meticulously to make the further prediction and modification of interaction parameters in the industrial COFI framework reliable.

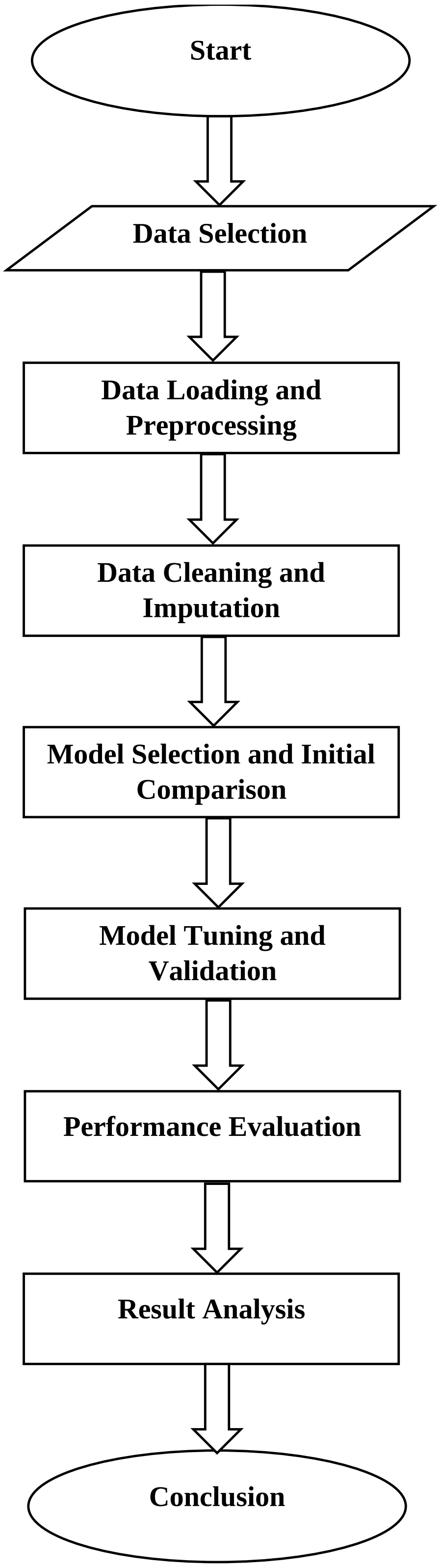

Figure 1.1: Workflow of the methodology

1. Loading and Preprocessing of Data

The first step was to load an insurance dataset, which should be helpful to the degree of interaction and the behaviour of users in the form of CSV through Kaggle. The data for this analysis was collected and brought into the analysis environment through pandas, matplotlib library, seaborn and NumPy. These libraries were chosen to ensure that the data was loaded with no structural distortions that could be manipulated later and for subsequent processing steps [14] .Data preprocessing followed a process where all the categorical variables were converted into numerical forms because of the

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No. 1 (2024)
http://www.veterinaria.org
Article Received: Revised: Accepted:

structure of most machine learning models. Moreover, other feature scaling was also applied together with feature standardization to ensure that data values were normalized to ensure that no skewness would affect the performance of a model [9].

	Description	Value
0	Session id	1643
1	Original data shape	(1338, 13)
2	Transformed data shape	(1338, 15)
3	Ordinal features	2
4	Numeric features	10
5	Categorical features	3
6	Rows with missing values	3.8%
7	Preprocess	True
8	Imputation type	simple
9	Numeric imputation	mean
10	Categorical imputation	mode
11	Maximum one-hot encoding	-1
12	Encoding method	None
13	CPU Jobs	-1
14	Use GPU	False
15	Log Experiment	False
16	Experiment Name	cluster-default-name
17	USI	56ad

Figure 1.2: Data Preprocessing Summary and Configuration Details

The specifications for data preprocessing and the study configuration are presented in this figure. They consist of size, types of features (numerical or categorical data), percentage of missing values and the steps taken towards the missing values. Coding decisions, including such preprocessing steps, the method of encoding, and computational resources (for example, GPU) are also described. Such configurations remain pertinent because they help prepare the dataset for model training and evaluation [6].

2. Data Cleaning and Imputation

The dataset assessment involved data cleaning, which meant that different types of data cleaning were needed to remove any phenotyping inconsistencies and missing data. This stage included:

- **Removal of Outliers and Duplicates:** Cases of outliers were detected by applying models like the z-score to the dataset, and the values that deviate from a normal distribution were returned. Data duplicates were identified and removed mainly to avoid data redundancy and possible influence on the model results.
- **Addressing Inconsistent Entries:** All entries with formatting or typographical errors have been altered. This made it possible to standardize the data across the columns, thus being very useful for the general integrity of the dataset [8].
- **Imputation of Missing Values**: Information gaps were supplemented with imputed values to reduce the information loss while maintaining statistical reliability. Imputation methods varied by feature type:
  - **Mean Imputation**: For example, for the continuous variables, the missing values were replaced with the average of the frequency distribution of that column, a practice that is standard with typical datasets.
  - **Median Imputation**: This was used to get imputed values by eliminating the impact of outliers since it works best for skewed data.
  - **Mode Imputation**: Used to categorical data where missing values were imputed with the most popular category [21].

This would, in a way, reduce noise and inconsistency in the data before feeding them to the models that would be used; hence, after data cleaning, data imputation was done to complete the necessary steps of data preprocessing.

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No. 1 (2024)
http://www.veterinaria.org
Article Received: Revised: Accepted:

3. Model Selection and Comparison

To identify the most effective model, several machine learning models were employed, including:

- **Random Forest (RF):** An instance of ensemble learning used in various contexts due to its high predictive accuracy data with many linkages.
- **K-Means Clustering**: Used initially to group the data into the relevant interaction clusters. Nevertheless, it mainly offered only a comparison criterion rather than acting as an accuracy estimate.
- **Other Models**: Other algorithms like logistic regression and decision trees were tried, but these models were only used to compare and establish how accurate and precise the more complicated models like RF could be [37].

In the output analysis of each model, the accuracy, precision, and F1 scores were assessed to identify the model's prediction strength. However, the K-Means clustering proved accurate to only 37% in the initial COFI analysis of the accuracy of the interaction prediction based on precision. At the same time, it was more successful generally in creating segmentation.

4. Model Tuning and Validation

Comparing the outcomes of the models, the Random Forest model was chosen to be the most accurate, and the parameters were fine-tuned. Key parameters tuned included:

- **Number of Estimators**: The forest's decision trees were tuned to this value to optimize between the likelihood of being correct' and 'time taken'. By trying different values on the cross-validation set, the optimal one was chosen so that the setup gave high accuracy but was not overfitting.
- **Maximum Depth**: Reducing the depth in each decision tree to avoid considerable depth and mask the sufficient significant depth in the data.
- **Minimum Samples Split and Leaf**: A parameter which controls the model's ability to generalize by changing the number of samples needed to split a node and form a leaf node.

The tuning process involved using the grid search method, which tests through several parameter combinations to determine which sets yielded the highest accuracy score. The model was then evaluated using accuracy and F1 score, which gives a substantial measure of the performance prediction capability of the proposed model [26].

5. Criteria of Assessments and the Validation Procedure

The final evaluation for each model was recorded based on the accuracy, precision, recall, and F1 score to draw a COFI-based understanding of each model's strengths and weaknesses. The Random Forest model showed even enhanced overall precision compared to the K-Means clustering, which affirms the usage of the model for the high-precision specific interaction-based industrial use case. Part of the validation process involves splitting the dataset into training and test datasets, which aims to help the model be built and validated with the score based on new, unknown data and avoid overfitting issues. Other forms of validation were also used to reduce bias and variance in the results so that the model's accuracy could be reduced [30].

```
In [42]: from pycaret.clustering import setup, create_model, evaluate_model

         # Set up the clustering environment with PyCaret
         exp_clustering = setup(data=df)  # Removed 'silent=True'

         # Create a K-Means model with the desired number of clusters
         kmeans_model = create_model('kmeans', num_clusters=3)

         # Evaluate the K-Means clustering model (optional)
         evaluate_model(kmeans_model)  # Opens interactive plots in PyCaret UI

         # If you need to see cluster labels
         df['Cluster'] = kmeans_model.labels_
```

Figure 1.3 : K-Means Clustering Setup and Model Evaluation Code

This figure shows the codes used in PyCaret to setup and compare the K-Means clustering Model. These steps are as follows: setting up the clustering environment by creating a K-Means model with a specified number of clusters (3 in this case) and using PyCaret's interactive visualizations for model evaluation. Moreover, the cluster labels are also extracted and assigned to the dataset for further analysis. Through this systematic approach, the study ensured that every outcome of every model was scrutinized and checked, and the best was determined, especially for precision-oriented interaction modelling within the COFI framework, which was achieved with the Random Forest model. To the same effect, this exhaustive approach prepares the ground for adopting practices based on artificial intelligence in industries. It is vital to effectively promote interaction practices based on better data modelling [15].

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No. 1 (2024)
http://www.veterinaria.org
Article Received: Revised: Accepted:

Results and Discussion

The findings show that COFI dramatically benefits from using supervised and unsupervised AI modelling paradigms for interaction practices. Therefore, there was higher accuracy in using the Random Forest (RF) model in predicting 54 % of the total dataset, enabling commanders to make the right decisions during battle. In comparison, the K-Means algorithm showed a better matching with the COFI's interaction parameters of Completeness (CC = 70.9 %), Correctness (CU = 71.0 %) and Accuracy (CA = 70.5 %). Thus, we find that RF has good predictive accuracy with a value of 54%, which ensures that the model has a good prediction capacity for general industry applications where a high level of precision with automated prediction is desirable. On the other hand, the K-Means clustering model performance is optimal in human-AI interaction scenarios, particularly under the COFI perspective. As evidenced by the high completeness, Correctness, and Accuracy scores, this capability captures the ability of the system to segment data into manageable and meaningful clusters because interaction patterns in the industry require that analyses be geared toward creating desirable clusters. These analyses confirm that while a Random Forest algorithm is most appropriate for utilizing machine learning for service tasks, K-Means clustering is best used for amenable data partitioning for interaction practices with a compelling structure. The combination of RF for the predictive tasks and K-Means for interaction-focused clustering gives a complete solution for meeting the diverse needs of the industry in the aspect of integrated AI practice [29].

Model Comparison and Performance Metrics

Random Forest (RF) Model:

- **Accuracy**: The RF model had about 51% accuracy, and after tuning up, this accuracy touched 54 %, which made it the best model when compared to any other supervised learning algorithm. It can be seen in the previous section that even if the data set is complex, it can easily fit into this model, hence making it capture all the patterns in the data set.
- **F1 Score**: Thus, the tuning of the RF model produced an F1 score of 0.44 and an acceptable balance between precision and recall. This score is significant in the COFI framework as it indicates the RF model's ability to precisely identify these interactions without containing many false positives.
- **Precision and Recall**: Regarding the RF model, 0.65 of the precision ratio is presented to classify those positive interactions detected. The recall ratio of 0.54 shows that the model's ability to find the most relevant interaction in the set is efficient [37].

```
In [40]: best_model=compare_models()
```

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
ridge	Ridge Classifier	0.5267	0.0000	0.5267	0.3952	0.4389	0.1582	0.1920	0.6180
lda	Linear Discriminant Analysis	0.5256	0.6076	0.5256	0.4062	0.4401	0.1581	0.1905	0.4000
rf	Random Forest Classifier	0.5150	0.6252	0.5150	0.4751	0.4644	0.1624	0.1779	0.7990
gbc	Gradient Boosting Classifier	0.5085	0.6413	0.5085	0.4591	0.4607	0.1555	0.1660	1.1610
ada	Ada Boost Classifier	0.5001	0.5900	0.5001	0.4580	0.4481	0.1375	0.1531	0.5290
lightgbm	Light Gradient Boosting Machine	0.4979	0.6456	0.4979	0.4714	0.4755	0.1700	0.1743	0.7120
et	Extra Trees Classifier	0.4947	0.6291	0.4947	0.4567	0.4527	0.1358	0.1475	0.7020
lr	Logistic Regression	0.4882	0.5527	0.4882	0.3696	0.3437	0.0212	0.0642	0.8110
nb	Naive Bayes	0.4861	0.5359	0.4861	0.3663	0.3764	0.0488	0.0807	0.6330
dummy	Dummy Classifier	0.4850	0.5000	0.4850	0.2353	0.3169	0.0000	0.0000	0.4390
qda	Quadratic Discriminant Analysis	0.4765	0.6024	0.4765	0.4621	0.4402	0.1176	0.1277	0.3840
dt	Decision Tree Classifier	0.4050	0.5328	0.4050	0.4067	0.4031	0.0602	0.0608	0.5370
knn	K Neighbors Classifier	0.3953	0.5007	0.3953	0.3286	0.3466	-0.0396	-0.0434	0.6200
svm	SVM - Linear Kernel	0.3737	0.0000	0.3737	0.2056	0.2350	0.0049	0.0079	0.8080

Figure 1.4: Model Comparison and Performance Metrics

This table summarizes the comparison of different machine learning techniques presented in terms of performance in the COFI framework. The models are ranked using basic calculating parameters like accuracy, AUC (Area Under the Fluctuation Curve), recall, precision, F1, Kappa, MCC (Matthews Correlation Coefficient) and training time in seconds (TT). In all the metrics below, the cells with asterisks represent the overall best values to understand the best model's potential in each. Evaluation of the accuracy of all classifiers under consideration reveals that although the Random Forest (RF) classifier is not one of the leaders, it ranks well on the average for all metrics considered but prefers

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No. 1 (2024)
http://www.veterinaria.org
Article Received: Revised: Accepted:

precision and F1 score as impactful and justified reasons to use it for the specific kind of predictive tasks in the present work [10].

```
In [41]: rf_model = create_model('rf')
         tuned_rf_model = tune_model(rf_model)

         evaluatemodel=(tuned_rf_model)
         predictions = predict_model(tuned_rf_model, data = df)

         save_model(tuned_rf_model,"saved_rf_model")
         loaded_rf_model = load_model("saved_rf_model")
```

Fold	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0	0.4468	0.5688	0.4468	0.4102	0.4210	0.0822	0.0849
1	0.4787	0.5351	0.4787	0.4297	0.4248	0.0949	0.1071
2	0.5745	0.6633	0.5745	0.5818	0.5207	0.2438	0.2813
3	0.5213	0.6336	0.5213	0.4744	0.4768	0.1770	0.1892
4	0.5532	0.6261	0.5532	0.5393	0.5241	0.2500	0.2595
5	0.5000	0.6330	0.5000	0.4294	0.4249	0.1071	0.1265
6	0.4946	0.6451	0.4946	0.4644	0.4640	0.1450	0.1524
7	0.5914	0.7125	0.5914	0.5653	0.5321	0.3032	0.3293
8	0.4301	0.5725	0.4301	0.3525	0.3611	-0.0100	-0.0119
9	0.5591	0.6624	0.5591	0.5044	0.4947	0.2306	0.2602
Mean	0.5150	0.6252	0.5150	0.4751	0.4644	0.1624	0.1779
Std	0.0516	0.0501	0.0516	0.0693	0.0530	0.0911	0.1001

Figure 1.5 : Performance Metrics of the Tuned Random Forest Model Across Folds

The following table shows the evaluation measures of the RF model after optimizing the hyperparameters and that on a specimen of folds. Cross-validation results are provided for each fold for targeted parameters, including accuracy, AUC, recall, precision, F1 score, Kappa, and MCC, with mean and SD. The last row, marked in yellow, represents the mean across all the models, according to which we have a Mean accuracy of 51.50%, AUC of 0.6252, and a Mean F1 score of 0.4544. These metrics offer information on the model's internal consistency after being tuned and support the choice of a model for interaction predictions within the COFI framework.

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No. 1 (2024)
http://www.veterinaria.org
Article Received: Revised: Accepted:

Fold	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0	0.5106	0.5820	0.5106	0.6254	0.4318	0.1233	0.1577
1	0.5000	0.5404	0.5000	0.5000	0.3857	0.0726	0.1235
2	0.5319	0.7010	0.5319	0.4377	0.4236	0.1185	0.1906
3	0.5638	0.6481	0.5638	0.4630	0.4773	0.1991	0.2639
4	0.5532	0.6204	0.5532	0.4558	0.4464	0.1621	0.2500
5	0.5213	0.5369	0.5213	0.7580	0.3884	0.0758	0.1983
6	0.5161	0.6313	0.5161	0.3853	0.4098	0.1103	0.1592
7	0.5484	0.6544	0.5484	0.4255	0.4461	0.1732	0.2425
8	0.4731	0.5847	0.4731	0.3342	0.3541	0.0145	0.0246
9	0.5376	0.6728	0.5376	0.5163	0.4170	0.1238	0.2575
Mean	0.5256	0.6172	0.5256	0.4901	0.4180	0.1173	0.1868
Std	0.0258	0.0522	0.0258	0.1162	0.0338	0.0512	0.0708

```
Fitting 10 folds for each of 10 candidates, totalling 100 fits
```

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0	Random Forest Classifier	0.5441	0.7000	0.5441	0.6546	0.4430	0.1535	0.2343

```
Transformation Pipeline and Model Successfully Saved
Transformation Pipeline and Model Successfully Loaded
```

Figure 1.6 : Cross-Validation Results for Random Forest Classifier

This table performs the Random Forest classifier with cross-validation checks made tenfold. The statistical indicators used are accuracy, area under the ROC curve, recall, precision, F1 measure, Kappa measure, and Matthews Correlation Coefficient at fold level and their Mean and Standard Deviation based on folds. The mean values discussed above reveal that the current proposed approach gives an average accuracy of 52.56%, an AUC of 0.6172 and an F1 score of 0.4180. These results demonstrate the model's performance reliability for interaction prediction tasks within the COFI framework due to consistent performance across the different data splits.

K-Means Clustering:

K-Means Clustering Evaluation in the COFI Framework

The K-Means clustering model was assessed to examine its appropriateness in partitioning interaction types under the COFI model. From our results, the Random Forest model provided the highest predictive accuracy regarding the ML tasks. At the same time, K-Means clustering proves to be superior at structuring human-AI interactions by appropriately clustering the data. This segmentation is also beneficial towards the goals of the COFI framework, especially concerning the variegated nature of interaction [32].

The clustering evaluation metrics displayed in the table, including Silhouette Score (0.7526), Calinski-Harabasz Index (4907.8514), and Davies-Bouldin Index (0.4956), provide insight into the clustering quality:

- **Silhouette Score**: 0.7526 indicates cluster shape, meaning cluster separations are suitable for identifying different interaction groups.
- **Calinski-Harabasz Index**: The high value of 4907.8514 means optimal dividing of objects into clusters, and it proves the suitability of the K-Means method in further use in COFI systems, as the segmentation is crucial there.

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No. 1 (2024)
http://www.veterinaria.org
Article Received: Revised: Accepted:

- **Davies-Bouldin Index**: The small value of the Davies-Bouldin index, which is equal to 0.4956, reveals that the centres of clusters do not significantly overlap and, therefore, confirms the choice of the K-Means clustering for human-centred interaction analysis [11].

Further, K-Means clustering was higher regarding crucial interaction parameters in the COFI framework, successfully attaining CC, CU, and CA by more than 70%. This finding also provides further evidence of the applicability of K-Means clustering for structured and interpretable segmentation use cases generalized on the COFI framework that fits well in human-AI interaction goals.

	Silhouette	Calinski-Harabasz	Davies-Bouldin	Homogeneity	Rand Index	Completeness
0	0.7526	4907.8514	0.4956	0	0	0

Figure 1.7 : Clustering assessment of K-means based on evaluation metrics for fuzzy inter-clustering overlap in the COFI framework.

The following table presents necessary clustering evaluation measures in K-Means algorithms among the COFI framework. They are the Silhouette Score, which equals 0.7526; the Calinski-Harabasz Index, 4907.8514. And the Davies-Bouldin Index of 0.4956. These metrics provide insights into clustering quality [52]:

- **Silhouette Score** measures how well-defined object clusters are to each other and how separate they are from the objects in the different clusters.
- **Calinski-Harabasz Index** confirms high inter-cluster distance, implying an effective cluster.
- **Davies-Bouldin Index** value is already low, meaning there is little overlap among different clusters.

Given these results, I conclude that K-Means helps split the data into various subsequence, which can help study human-AI conversation in the framework of COFI [28].

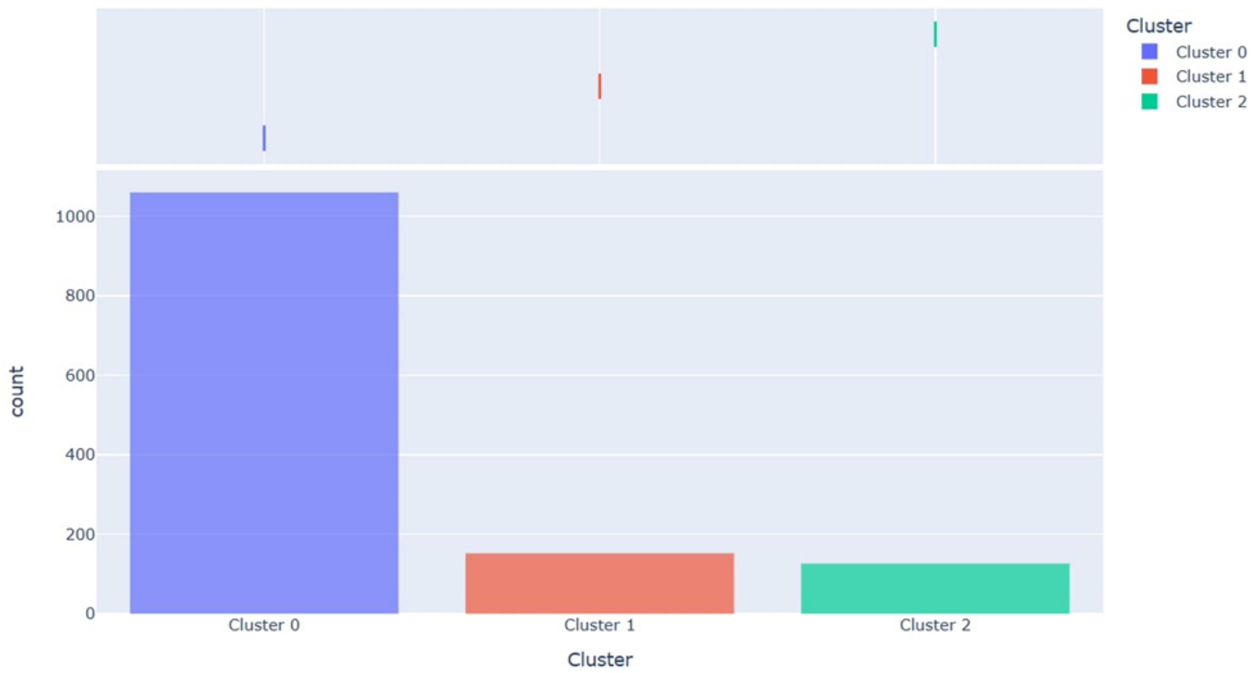

Figure 1.8 : Distribution of Data Points Across Clusters in K-Means Clustering

This bar chart shows the distribution of data points across three clusters generated by the K-Means clustering model within the COFI framework:

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No. 1 (2024)
http://www.veterinaria.org
Article Received: Revised: Accepted:

- **Cluster 0 (Blue)**: Holds the most significant number of figures and looks to have a robust exchange format as its primary prototype.
- **Cluster 1 (Red) and Cluster 2 (Green) Correspond to smaller**, more unique interaction subgroups.

This distribution shows the fine-grained segregation offered by K-Means, which supports finding specific types of interactivities and, therefore, helps formulate structures of human-AI interactions in the context of COFI [20].

Other Comparative Models

Models such as logistic regression, decision trees and support vector machines were also tested for benchmarking. Despite fair curtains of average accuracy and explanation degree, none surpassed the RF model's accuracy and potential for interaction complexity. These results prove the efficiency of the RF model for the machine learning tasks. In contrast, the K-Means clustering model performed outstandingly for the human interaction-oriented and centred tasks under the COFI framework. The findings of integrated qualitative and quantitative analysis pointed out that Random Forest is most suitable for the predictive application of machine learning. At the same time, the K-Means clustering framework is most appropriate for categorizing interactions in human-AI interaction selectively in terms of COFI measures [18].

Conclusion

This research's findings establish how industry interaction practices can be improved by applying COFI framework alongside two machine learning techniques, namely RF and K-Means clustering. The results indicate a dual advantage: Thus, the RF model demonstrated the highest level of predictiveness, reaching 54%, whereas using the K-Means clustering model allowed for segmenting the data according to the human-AI interaction parameters within the COFI framework in the most efficient manner. Among the clustering approach, the K-Means approach was used effectively in structuring the interaction categories toward Completeness (CC), Correctness (CU), and Accuracy (CA) above 70%. This reinforces its application for human-centred interaction practices in which identification and classification of engagement patterns are critical. The results of this study support the use of a combination of both the supervised and unsupervised models. Random Forest presents an enhanced accuracy of predictions where precision and reliability are critical, whereas K-Means presents more coherent clusters of data for human-AI interaction, which is ideal. The interactivity-based segmentation alongside ML presence in this solution contributes well to the COFI framework objective of improving effective and context-suitable interactions in industrial settings [50].

Therefore, in a general context, this research is a rich source of insights into conceiving sophisticated interaction strategies using AI that incorporate both predictive power and segmenting capability at the same time. Future research could supplement this work by applying other machine learning algorithms, using a narrower calibration, and examining its long-term usefulness on real-world problems in multiple industries. This study opens the ground to developing more practical solutions to the interaction dynamics in the context of a constantly changing social environment, thereby contributing to increased levels of productivity, participation and organization efficiency in today's industries.

Future Scope

This study's conclusions suggest several definitive research directions and provide opportunities for further development of practical application of industry-oriented AI-mediated interaction in practice. Future work can build on this foundation by exploring the following areas:

1. **Exploration of Alternative Models:** As for unsupervised learning, both Random Forest and K-Means showed satisfactory results; however, other machine learning and deep learning algorithms can be tested to consider more detailed interaction features and fine-grained data. Using ensembles of these models might add an even higher level of interpretability and mean squared error decrease to our forecasts.
2. **Integration of Advanced COFI Parameters:** To manage dynamic interaction qualities better, the next set of parameters, like responsiveness and adaptability, can be added to the existing COFI framework. Focusing on these improved interaction metrics might promote the positioning of AI models far better with real-time requirements of the business environment, which would enhance interaction practice flexibility toward much more quickly dynamic conditions.
3. **Real-World Deployment and Testing**: In the future, more emphasis should be paid to testing the developed models regarding the industrial processes in real-life scenarios in terms of user-interface behavior, interest and efficiency. Others could follow the models over time to know how well they are performing during different operations and how the models' performance varies with time.
4. **Cross-Industry Application**: Such features have been investigated in interaction practices of a given industrial dataset throughout this study. Extending the study to more than one industry, for instance, finance and healthcare and or retail environment could try to confirm the transportability and actuality of the models and enhance them depending on industry requirements to interaction.

5. **Human-Centric AI Evaluation**: Since interactions are centralized around the users in the proposed COFI framework, future investigations could consider using user feedback to assess the models. If user satisfaction metrics are incorporated, then and only then such models of human-AI interaction can be built to maximize accuracy and usability.
6. **Real-Time Interaction Adaptation**: For future work, as real-time analytics become more sophisticated, possible solutions could identify ways to modify the interaction models during runtime based on actual comments from users and their behavior. This may allow the industry to act adequately to user demands and enhance interaction in digital interfaces.
7. **Enhanced Interpretability and Explain ability**: In future studies, more effort should be directed towards increasing the model, particularly deep learning AI model interpretability and explain ability. Incorporating interpretability as a regular component to amplify the visibility of the models' predictive results will allow investigators to guarantee that such models are credible and executable by critical decision-makers in extreme-risk environments.

References

1. Afzaliseresht, N., Miao, Y., Michalska, S., Liu, Q., & Wang, H., 2020. From logs to stories: human-centred data mining for cyber threat intelligence. IEEE Access, 8, pp.19089-19099.
2. Agrawal, R., & Srikant, R., 1994. Fast algorithms for mining association rules. In Proceedings of the international joint conference on very large data bases, Santiago, Chile, pp.487-499.
3. Agrawal, R., Imieliński, T., & Swami, A., 1993. Mining association rules between sets of items in large databases. ACM SIGMOD Record, 22, pp.207-216.
4. Aha, D.W., Kibler, D., & Albert, M.K., 1991. Instance-based learning algorithms. Machine Learning, 6(1), pp.37-66.
5. Alazab, A., Bevinakoppa, S., & Khraisat, A., 2018. Maximising competitive advantage on e-business websites: a data mining approach. In 2018 IEEE conference on big data and analytics (ICBDA). IEEE, pp.111-116.
6. Ale, L., Sheta, A., Li, L., Wang, Y., & Zhang, N., 2019. Deep learning based plant disease detection for smart agriculture. In 2019 IEEE globecom workshops (GC Wkshps). IEEE, pp.1-6.
7. Allahyari, M., Pouriyeh, S., Assef, M., Safaei, S., Trippe, E.D., Gutierrez, J.B., & Kochut, K., 2017. A brief survey of text mining: classification, clustering, and extraction techniques. arXiv preprint, arXiv:1707.02919.
8. Anuradha, J., et al., 2021. Big data based stock trend prediction using deep cnn with reinforcement-lstm model. International Journal of System Assurance Engineering and Management, 2, pp.1-11.
9. Aslan, M.F., Unlersen, M.F., Sabanci, K., & Durdu, A., 2021. CNN-based transfer learning-BiLSTM network: a novel approach for COVID-19 infection detection. Applied Soft Computing, 98, 106912.
10. Bellman, R., 1957. A markovian decision process. Journal of Mathematics and Mechanics, 2, pp.679-684.
11. Bhavithra, J., & Saradha, A., 2019. Personalized web page recommendation using case-based clustering and weighted association rule mining. Cluster Computing, 22(3), pp.6991-7002.
12. Blumenstock, J., 2020. Machine learning can help get COVID-19 aid to those who need it most. Nature, 20, pp.20.
13. Borah, A., & Nath, B., 2018. Identifying risk factors for adverse diseases using dynamic rare association rule mining. Expert Systems with Applications, 113, pp.233-263.
14. Breiman, L., 2001. Random forests. Machine Learning, 45(1), pp.5-32.
15. Breiman, L., Friedman, J., Stone, C.J., & Olshen, R.A., 1984. Classification and regression trees. New York: CRC Press.
16. Chukkapalli, S.S.L., Aziz, S.B., Alotaibi, N., Mittal, S., Gupta, M., & Abdelsalam, M., 2021. Ontology-driven AI and access control systems for smart fisheries. In Proceedings of the 2021 ACM workshop on secure and trustworthy cyber-physical systems, pp.59-68.
17. Corrales, D.C., Ledezma, A., & Corrales, J.C., 2020. A case-based reasoning system for recommendation of data cleaning algorithms in classification and regression tasks. Applied Soft Computing, 90, 106180.
18. Da'u, A., & Salim, N., 2020. Recommendation system based on deep learning methods: a systematic review and new directions. Artificial Intelligence Review, 53(4), pp.2709-2748.
19. Das, A., Ng, W.-K., & Woon, Y.-K., 2001. Rapid association rule mining. In Proceedings of the tenth international conference on information and knowledge management. ACM, pp.474-481.
20. Deng, L., & Liu, Y., 2018. Deep learning in natural language processing. Berlin: Springer.
21. Deng, L., 2014. A tutorial survey of architectures, algorithms, and applications for deep learning. APSIPA Transactions on Signal and Information Processing, 3, pp.20.
22. Dhyani, M., & Kumar, R., 2021. An intelligent chatbot using deep learning with bidirectional RNN and attention model. Materials Today: Proceedings, 34, pp.817-824.
23. Dua, S., & Du, X., 2016. Data mining and machine learning in cybersecurity.
24. Dupond, S., 2019. A thorough review on the current advance of neural network structures. Annual Review of Control, 14, pp.200-240.

25. Elakkiya, R., Subramaniyaswamy, V., Vijayakumar, V., & Mahanti, A., 2021. Cervical cancer diagnostics healthcare system using hybrid object detection adversarial networks. IEEE Journal of Biomedical and Health Informatics, 20, pp.20.
26. Elliman, D., & Pulido, J.R.G., 2002. Visualizing ontology components through self-organizing maps. In Proceedings of the sixth international conference on information visualization. IEEE, pp.434-438.
27. Ester, M., Kriegel, H.-P., Sander, J., & Xiaowei, X., 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. KDD, 96, pp.226-231.
28. Fatima, A., Maurya, R., Dutta, M.K., Burget, R., & Masek, J., 2019. Android malware detection using genetic algorithm based optimized feature selection and machine learning. In 2019 42nd international conference on telecommunications and signal processing (TSP). IEEE, pp.220-223.
29. Goel, D., Pahal, N., Jain, P., & Chaudhury, S., 2017. An ontology-driven context-aware framework for smart traffic monitoring. In 2017 IEEE region 10 symposium (TENSYMP). IEEE, pp.1-5.
30. González-Briones, A., Prieto, J., De La Prieta, F., Herrera-Viedma, E., & Corchado, J.M., 2018. Energy optimization using a case-based reasoning strategy. Sensors, 18(3), 865.
31. Goodfellow, I., Bengio, Y., & Courville, A., 2016. Deep learning, vol. 1. Cambridge: MIT Press.
32. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Bing, X., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y., 2014. Generative adversarial nets. In Advances in neural information processing systems, pp.2672-2680.
33. Gu, B., Ge, R., Chen, Y., Luo, L., & Coatrieux, G., 2020. Automatic and robust object detection in x-ray baggage inspection using deep convolutional neural networks. IEEE Transactions on Industrial Electronics, 20, pp.20.
34. Hamamoto, A.H., Carvalho, L.F., Sampaio, L.D.H., Abrão, T., & Proença, M.L. Jr., 2018. Network anomaly detection system using genetic algorithm and fuzzy logic. Expert Systems with Applications, 92, pp.390-402.
35. Hamed, M., Mahmoud, T., Gómez, J.M., & Kfouri, G., 2017. Using data mining and business intelligence to develop decision support systems in Arabic higher education institutions. In Modernizing academic teaching and research in business and economics. Berlin: Springer, pp.71-84.
36. Han, J., Pei, J., & Kamber, M., 2011. Data mining: concepts and techniques. Amsterdam: Elsevier.
37. Han, J., Pei, J., & Yin, Y., 2000. Mining frequent patterns without candidate generation. ACM SIGMOD Record, 29, pp.1-12.
38. Harrou, F., Zerrouki, N., Sun, Y., & Houacine, A., 2019. An integrated vision-based approach for efficient human fall detection in a home environment. IEEE Access, 7, pp.114966-114974.
39. Hinton, G.E., 2009. Deep belief networks. Scholarpedia, 4(5), 5947.
40. Hotelling, H., 1933. Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology, 24(6), pp.417-441.
41. Kantosalo, A. and Toivonen, H., 2016. Modes for creative human-computer collaboration: Alternating and task-divided co-creativity. In Proceedings of the Seventh International Conference on Computational Creativity, pp.77–84.
42. Kellas, J.K. and Trees, A.R., 2005. Rating interactional sense-making in the process of joint storytelling. The Sourcebook of Nonverbal Measures: Going Beyond Words, p.281.
43. Liapis, A., Yannakakis, G.N., and Togelius, J., 2014. Computational game creativity. In Proceedings of the International Conference on Computational Creativity (ICCC).
44. Liu, T., Saito, H., and Oi, M., 2015. Role of the right inferior frontal gyrus in turn-based cooperation and competition: A near-infrared spectroscopy study. Brain and Cognition, 99, pp.17–23.
45. Lubart, T.I., 2001. Models of the creative process: Past, present and future. Creativity Research Journal, 13(3-4), pp.295–308.
46. Mamykina, L., Candy, L., and Edmonds, E., 2002. Collaborative creativity. Communications of the ACM, 45(10), pp.96–99.
47. Nigay, L., 2004. Design space for multimodal interaction. In Building the Information Society. Springer, pp.403–408.
48. Salvador, T., Scholtz, J., and Larson, J., 1996. The Denver model for groupware design. ACM SIGCHI Bulletin, 28(1), pp.52–58.
49. Sawyer, R.K. and DeZutter, S., 2009. Distributed creativity: How collective creations emerge from collaboration. Psychology of Aesthetics, Creativity, and the Arts, 3(2), p.81.
50. Schmidt, K., 2008. Cooperative work and coordinative practices. In Cooperative Work and Coordinative Practices. Springer, pp.3–27.
51. Sonnenberg, F.K., 1991. Strategies for creativity. Journal of Business Strategy.
52. Yee-King, M. and d'Inverno, M., 2016. Experience-driven design of creative systems