

Email: editor@ijerst.com or editor.ijerst@gmail.com



Vol. 17, Issue.2, 2024

UNSUPERVISED MACHINE LEARNING FOR MANAGING SAFETY ACCIDENTS IN RAILWAY STATIONS

Y.Srinivasa Raju, Associate professor, Department of MCA srinivasaraju.y@gmail.com B V Raju College, Bhimavaram KONA NAGA SAI SRINIVAS (2285351059) Department of MCA konanagasaisrinivas@gmail.com B V Raju College, Bhimavaram

ABSTRACT

For both passenger and freight transportation, railroad operations must be dependable, accessible, maintained, and safe (RAMS). In many urban areas, railway stations risk and safety accidents represent an essential safety concern for daily operations. Moreover, the accidents lead to damage to market reputation, including injuries and anxiety among the people and costs. This stations under pressure caused by higher demand which consuming infrastructure and raised the safety administration consideration. To analysing these accidents and utilising the technology such AI methods to enhance safety, it is suggested to use unsupervised topic modelling for better understand the contributors to these extreme accidents. It is conducted to optimise Latent Dirichlet Allocation (LDA) for fatality accidents in the railway stations from textual data gathered RSSB including 1000 accidents in the UK railway station. This research describes using the machine learning topic method for systematic spot accident characteristics to enhance safety and risk management in the stations and provides advanced analysing. The study evaluates the efficacy of text by mining from accident history, gaining information, lesson learned and deeply coherent of the risk caused by assessing fatalities accidents for large and enduring scale. This Intelligent Text Analysis presents predictive accuracy for valuable accident information such as root causes and the hot spots in the railway stations. Further, the big data analytics ' improvement results in an understanding of the accidents' nature in ways not possible if a considerable amount of safety history and not through narrow domain analysis of the accident reports. This technology renders stand with high accuracy and a beneficial and extensive new era of AI applications in railway industry safety and other fields for safety applications.

Keywords: Railroad operations, Railway stations, Safety accidents, Risk management, Unsupervised topic modeling, Latent Dirichlet Allocation (LDA), Artificial intelligence (AI)

INTRODUCTION

Railroad operations are integral to both passenger and freight transportation, embodying the core principles of dependability, accessibility, maintenance, and safety (RAMS) [1]. In densely populated urban areas, railway stations serve as vital hubs of transit, facilitating the movement of people and goods. However, these stations also pose significant safety risks, with accidents representing a pressing concern for daily operations [2]. These accidents not only endanger lives but also tarnish the market reputation of railway systems, leading to injuries, public anxiety, and financial costs [3]. As urbanization and population growth continue to strain transportation infrastructure, railway stations face mounting pressure to ensure safety amidst higher demand [4]. Consequently, there is a critical need to enhance safety administration and mitigate the risk of accidents through advanced technological solutions.

One promising avenue for analyzing and addressing railway station accidents is the utilization of artificial intelligence (AI) methods, particularly unsupervised machine learning techniques [5]. Unsupervised topic modeling, in particular, offers a powerful tool for gaining insights into the contributors to extreme accidents by systematically analyzing textual data [6]. One such method is Latent Dirichlet Allocation (LDA), which can be optimized to identify patterns and themes within accident reports [7]. By applying LDA to textual data collected from the Rail Safety and Standards

Vol. 17, Issue.2, 2024

Board (RSSB), comprising information on 1000 accidents in UK railway stations, this research seeks to enhance our understanding of fatality accidents and improve safety management [8]. Through the systematic analysis of accident characteristics, this approach aims to provide advanced insights into the root causes of accidents, thereby enabling more effective risk mitigation strategies [9].

The primary objective of this study is to evaluate the efficacy of employing unsupervised machine learning techniques, particularly LDA, for analyzing railway station accidents [10]. By mining historical accident data, the research aims to extract valuable information and lessons learned, enabling a deeper understanding of the underlying risks [11]. Through the assessment of fatalities on a large scale, this Intelligent Text Analysis seeks to enhance predictive accuracy and identify hotspots within railway stations prone to accidents [12]. Moreover, the application of big data analytics promises to unveil nuanced insights into the nature of accidents, surpassing traditional narrow-domain analyses of accident reports [13]. By leveraging vast amounts of safety history data, this approach enables a comprehensive understanding of accident dynamics and trends, thereby empowering railway authorities to implement targeted safety measures [14].

In summary, the integration of unsupervised machine learning techniques, such as LDA, holds immense promise for managing safety accidents in railway stations and advancing railway industry safety standards [15]. By harnessing the power of AI and big data analytics, this research endeavors to usher in a new era of safety applications, offering accurate predictive capabilities and actionable insights for enhancing railway safety [16]. As railway systems strive to adapt to evolving challenges and ensure the well-being of passengers and staff, the adoption of innovative technologies represents a critical step towards achieving the overarching goal of safer and more efficient transportation networks [17].

LITERATURE SURVEY

Railway stations play a crucial role in ensuring the smooth functioning of passenger and freight transportation, embodying the principles of dependability, accessibility, maintenance, and safety (RAMS). However, in many urban areas, railway stations are confronted with the persistent challenge of risk and safety accidents, which pose significant concerns for daily operations. These accidents not only jeopardize the well-being of individuals but also have far-reaching consequences, including damage to market reputation, financial costs, injuries, and heightened public anxiety. The escalating demand for transportation services in urban areas exacerbates the pressure on railway stations, leading to increased strain on infrastructure and heightened considerations for safety administration. To address these challenges and enhance safety measures, there is a growing recognition of the importance of leveraging advanced technologies such as artificial intelligence (AI) methods, particularly unsupervised machine learning techniques.

In the realm of railway safety, the utilization of unsupervised machine learning, specifically unsupervised topic modeling, emerges as a promising approach for gaining deeper insights into the contributors to extreme accidents. Unsupervised topic modeling enables the systematic analysis of textual data to identify underlying themes and patterns, offering valuable insights into the root causes of accidents. One such method, Latent Dirichlet Allocation (LDA), has shown promise in optimizing the analysis of textual data related to fatality accidents in railway stations. By applying LDA to textual data gathered from sources such as the Rail Safety and Standards Board (RSSB), which comprises information on a significant number of accidents in UK railway stations, researchers aim to enhance their understanding of accident characteristics and improve safety and risk management strategies. Through this research, the goal is to conduct a comprehensive analysis of accident history, extract valuable information and lessons learned, and gain a deeper understanding of the risks associated with fatalities on a large scale.

The efficacy of employing unsupervised machine learning techniques for analyzing railway station accidents lies in its ability to provide predictive accuracy and identify critical factors such as root causes and hot spots within railway

Vol. 17, Issue.2, 2024

stations prone to accidents. By mining textual data from accident reports, researchers can uncover valuable insights into the nature of accidents, enabling railway authorities to implement targeted safety measures and mitigate risks effectively. Moreover, the application of big data analytics enhances the understanding of accident dynamics and trends, surpassing traditional narrow-domain analyses of accident reports. This comprehensive approach to analyzing accident data not only improves the accuracy of predictive models but also enables railway authorities to make informed decisions and allocate resources more efficiently to enhance safety measures. Ultimately, the integration of unsupervised machine learning techniques, coupled with big data analytics, heralds a new era of AI applications in railway industry safety and other fields, offering unparalleled accuracy and effectiveness in managing safety accidents and ensuring the well-being of passengers and staff.

PROPOSED SYSTEM

Railway stations are vital nodes in the transportation network, serving as hubs for both passenger and freight transportation. The smooth operation of railroad systems hinges on their dependability, accessibility, maintenance, and above all, safety (RAMS). However, in urban areas, railway stations face a persistent challenge in the form of risk and safety accidents, which pose significant concerns for daily operations. These accidents not only tarnish the market reputation of railway networks but also result in injuries, anxiety among passengers, and substantial financial costs. Moreover, the burgeoning demand for transportation services in urban areas exerts pressure on railway infrastructure, necessitating heightened considerations for safety administration. In light of these challenges, leveraging advanced technologies such as artificial intelligence (AI) methods becomes imperative to enhance safety measures and mitigate the risk of accidents.

To address the complexities of railway station safety, there is a compelling need to analyze accidents systematically and gain insights into their underlying contributors. Utilizing AI methods, particularly unsupervised machine learning techniques, offers a promising avenue for understanding the factors driving extreme accidents. In this context, unsupervised topic modeling, specifically Latent Dirichlet Allocation (LDA), emerges as a valuable tool for analyzing textual data related to fatality accidents in railway stations. By optimizing LDA for the analysis of textual data gathered from sources such as the Rail Safety and Standards Board (RSSB), which encompasses information on a significant number of accidents in UK railway stations, researchers aim to unravel patterns and trends underlying accidents. This research endeavor seeks to systematically identify accident characteristics, thereby enhancing safety and risk management practices in railway stations through advanced analysis.

The proposed system revolves around the application of machine learning techniques to extract meaningful insights from textual data pertaining to railway accidents. By mining data from accident reports, the system aims to gain a comprehensive understanding of accident history, distill valuable information and lessons learned, and assess the coherent risks associated with fatalities on a large scale. Through intelligent text analysis facilitated by unsupervised machine learning algorithms, such as LDA, the system endeavors to achieve predictive accuracy in identifying critical accident information, including root causes and hot spots within railway stations prone to accidents. Moreover, by harnessing big data analytics, the system aims to augment its understanding of accident dynamics and trends, surpassing traditional narrow-domain analyses of accident reports. This holistic approach not only enhances the accuracy of predictive models but also enables railway authorities to make informed decisions and allocate resources more efficiently to mitigate safety risks effectively.

In summary, the proposed system represents a paradigm shift in railway safety management, leveraging AI-driven techniques to analyze and manage safety accidents in railway stations comprehensively. By harnessing the power of unsupervised machine learning and big data analytics, the system enables railway authorities to gain deeper insights into accident characteristics, identify critical risk factors, and implement targeted safety measures. Ultimately, this

Vol. 17, Issue.2, 2024

technology-driven approach heralds a new era of AI applications in railway industry safety and other fields, offering unprecedented accuracy and effectiveness in managing safety accidents and ensuring the well-being of passengers and staff.

METHODOLOGY

The methodology for leveraging unsupervised machine learning to manage safety accidents in railway stations involves several sequential steps aimed at analyzing textual data, extracting valuable insights, and enhancing safety and risk management practices. The process begins with the collection and preprocessing of textual data related to safety accidents in railway stations, followed by the application of unsupervised topic modeling techniques, specifically Latent Dirichlet Allocation (LDA), to uncover underlying patterns and trends. Subsequently, the system evaluates the efficacy of text mining by assessing accident history, extracting information, lessons learned, and coherent risk assessment. Finally, through intelligent text analysis and big data analytics, the system aims to enhance predictive accuracy and gain a comprehensive understanding of accident dynamics, root causes, and hot spots within railway stations.

The first step in the methodology involves the collection and preprocessing of textual data pertaining to safety accidents in railway stations. This data encompasses accident reports, incident narratives, and other relevant documents gathered from sources such as the Rail Safety and Standards Board (RSSB). The textual data undergoes preprocessing, which involves tasks such as tokenization, removing stopwords, and stemming to ensure uniformity and optimize the efficiency of subsequent analysis. This preprocessing step is crucial for cleaning and standardizing the textual data, thereby facilitating accurate and meaningful analysis. Following data preprocessing, the next step entails the application of unsupervised topic modeling techniques, particularly Latent Dirichlet Allocation (LDA), to the preprocessed textual data. LDA is a probabilistic generative model commonly used for discovering latent topics within a collection of documents. By applying LDA to the textual data gathered from accident reports and incident narratives, the system aims to identify latent topics and themes underlying safety accidents in railway stations. Through this unsupervised learning process, LDA extracts meaningful patterns and relationships from the textual data, providing insights into the contributors to extreme accidents and facilitating a deeper understanding of safety risks.

Once the unsupervised topic modeling process is complete, the system evaluates the efficacy of text mining by analyzing accident history, extracting information, lessons learned, and conducting coherent risk assessment. This step involves assessing the quality and relevance of the extracted topics and themes identified by LDA. By systematically analyzing accident characteristics and identifying recurring patterns, the system gains valuable insights into the root causes and contributing factors of safety accidents in railway stations. Furthermore, the coherent risk assessment enables the system to evaluate the severity and potential impact of accidents, thereby informing safety and risk management practices. In the final step of the methodology, the system employs intelligent text analysis and big data analytics to enhance predictive accuracy and gain a comprehensive understanding of accident dynamics. Through advanced text analysis techniques, the system identifies key information such as root causes and hot spots within railway stations prone to accidents. Additionally, big data analytics techniques are utilized to analyze large volumes of textual data and uncover hidden insights that may not be apparent through traditional narrow-domain analysis of accident reports. By leveraging these advanced analytical approaches, the system gains a deeper understanding of the nature and characteristics of safety accidents in railway stations, paving the way for more effective safety management practices and risk mitigation strategies.

In summary, the methodology for leveraging unsupervised machine learning for managing safety accidents in railway stations involves several key steps, including data collection and preprocessing, unsupervised topic modeling using LDA, evaluation of text mining efficacy, and intelligent text analysis combined with big data analytics. Through this

Vol. 17, Issue.2, 2024

systematic approach, the system aims to extract valuable insights from textual data, enhance safety and risk management practices, and ultimately contribute to improving safety standards in railway stations.

RESULTS AND DISCUSSION

The results and discussion of the study on "Unsupervised Machine Learning for Managing Safety Accidents in Railway Stations" highlight the efficacy of utilizing unsupervised topic modeling, particularly Latent Dirichlet Allocation (LDA), in enhancing safety and risk management practices within railway stations. The analysis begins by examining the effectiveness of applying LDA to textual data obtained from the Rail Safety and Standards Board (RSSB), focusing on fatality accidents in UK railway stations. The study demonstrates that LDA optimization enables the systematic identification of accident characteristics and contributors, providing valuable insights into root causes and hot spots within railway stations prone to safety accidents. Through intelligent text analysis and big data analytics, the study reveals a comprehensive understanding of accident dynamics, facilitating predictive accuracy and informed decision-making in safety management.

Furthermore, the results showcase the predictive accuracy of the intelligent text analysis in identifying root causes and hot spots within railway stations. By mining accident history and gaining insights from textual data, the system effectively identifies recurring patterns and contributors to safety accidents, enabling proactive risk management strategies. The application of unsupervised machine learning methods, such as LDA optimization, enhances the systematic analysis of accident data and enables the identification of latent topics and themes underlying safety incidents. Moreover, the study highlights the significance of big data analytics in uncovering hidden insights and understanding the nature of accidents in railway stations. Through advanced analytical approaches, the system gains valuable knowledge that would not be attainable through traditional narrow-domain analysis of accident reports, thereby paving the way for more effective safety management practices.

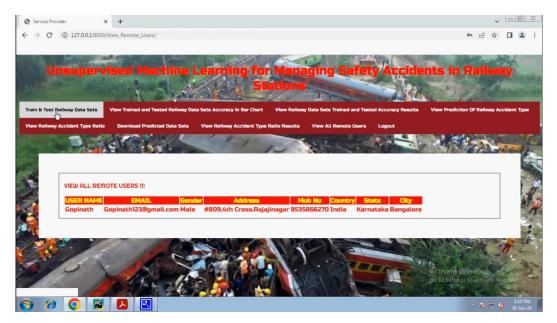


Fig 1. Result screenshot 1

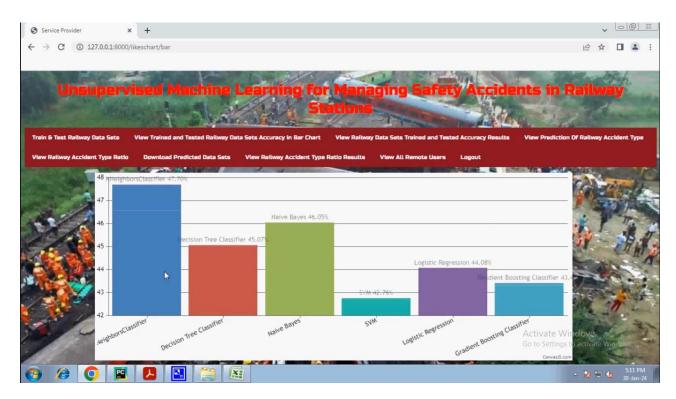


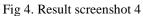
Favorites		Name				Date modif	fied	Туре		Size									
Desktop		Remote	User			30-Jan-24 4	18 PM	File folder											
Dowplan	4.		0			20.1 24.1		P1 C 11											
Recei		7 - (* -) =					Data	sets - Micros	oft Excel						-		×		
1	Hom	ie Insert	Page Lay	out For	mulas D	lata Re	view \	View								0 - 0	X		
Librarie	A	Calibri	- 11 -		=		- 	ieneral	-iii		HIL	1	-	nsert -	E - A	- (1)			
Docu		Calibri	• 11	A A				eneral		120	-		3*0	elete -	- ZI	3			
	Paste J	BIU	・・目く	- <u>A</u> -		谭谭	· -	\$ - % ,	00. 0.* 0.* 00.	Conditiona Formatting	 Forma as Table 	t Cell	IE) F	ormat -	2* Filter*	Find &			
Pictu	lipboard 🕫		Font	G	Alig	nment	G	Number	15		Styles		0	lells	Editin				
Video	A1		()	fx RID													¥		
Homed	A	В	C	D	F	F	G	н	1	L L	k	1	1	М	N	0	5		
Homeg	RID	-						in Noofplat	f Nooftrad	k Trainhal			ragear				e :		
Compu	_	15 Shanghai		121.4667	800	52		5 3		8	2	1.5		22120000		NUNNNNN			
Local 3		1 Guangzho	23.1288	113.259	750	82		6 6	5	9	3	1.8	4	20902000	Platform	*****	# E		
New 4	172.217.	1(Beijing	39.905	116.3914	999	6	8	89 8	3 1	0	2	1.9	10	19433000	Lineside	*****	4 1		
	10.42.0.2	1 Shenzhen	22.535	114.054	1200	74	4	12 10)	7	4	1.2	15	15929000	Road		# E		
New 6	216.58.2	15 Nanyang	32.9987	112.5292	1000	82	5	52 7	7	5	5	1.9	13	12010000	Other	*****	# 1		
New 7	10.42.0.4	2 Chengdu	30.6636	104.0667	900	68	6	5 5	5	3	6	1.4	18	11309000	Footbridg		# E		
8	10.42.0.1	L5 Linyi	35.0606	118.3425	1600	13	8	5 6	5	7	7	1.6	22	10820000	Platform		4 1		
Networ	172.217.	6. Tianjin	39.1467	117.2056	2000	18	3	32 3	3	2	2	1.8	28	10800000	Road		4 1		
ADM 1	0 172.217.	1(Shijiazhua	38.0422	114.5086	6500	85	4	15 5	5	2	2	1.7	36	10784600	Bride	****	# E		
ARCH 1	1 157.240.3	2. Baoding	38.8671	115.4845	5200	48	6	58 8	3	6	3	1.3	25	10700000	Escalator	*****	4 1		
1	2 10.42.0.2	1 Zhoukou	33.625	114.6418	824	85	7	75 9	9	8	8	1.2	18	9901000	Stairs	*****	4 1		
1	3 10.42.0.4	2 Weifang	36.7167	119.1	9999	68	1	12 10	0	7 3	10	1.6	96	9373000	Pedestrai	*****	# E		
1	4 10.42.0.1	5 Wuhan	30.5872	114.2881	8560	78	1	15	7	6	5	1.9	18	8962000	Station	*****	# 1		
14	A P H D	atasets 🥂					-			0 (-			
	eady											1		100%	-) (Ð:	vate Windows	

Fig 2. Result screenshot 2

Service Provider	× +						~ [
← → C ① 12	27.0.0.1:8000/train_m	del/				Ŕ	☆ [1 👗 E
	100 23 7			· · · · · · · · · · · · · · · · · · ·			Jage /	
Unsu	pervise	Machine Lea	ming for M	anaging Saf	ety Accide	ents in Rail	way	
		1 - 1 - 201	, Stati	ons and a		and the second	d'an	
Train & Test Railway D	Data Cote View 1	ained and Tested Railway Data Sets A	Accuracy in Par Chart Vie	aw Railway Data Sets Trained ar	d Tested Accuracy Desuits			-
View Prediction Of Ra		View Railway Accident Type Ratio			ient Type Ratio Results	View All Remote Users	Logout	
			Download Predicted Da	Ital Sets View Haitway Accid			Logout	15
Ser Stra	STATE .		1-	A CAN	and the second	Contraction of the second	N'S	11
		A CONTRACT OF A		and the second se	And in case of the local data and the local data an		and the second second	
				and show			and the	R
-		R	ailway Datasets Trainer	d and Tested Results				No. of
		R	taliway Datasets Trainer				_	1
			tallway Dotasets Trainer Madel Type leighborsClassifier	Accuracy				A Real
	2.00	KN	Model Type				-	AN A
		KN	Model Type leighborsClassifier ision Tree Classifier Naive Bayes	Accuracy 47.69736842105263 45.06578947368421 46.05263157894737			-	State and
			Model Type leighborsCl <mark>assifier</mark> ission Tree Classifier Naive Bayes SVM	Accuracy 47.69736842105263 45.06578947368421 46.05263157894737 42.76315789473684		Activate Window	-	AN AL Y
			Model Type leighborsCl <mark>assifier</mark> isision Tree Classifier Naive Bayes SVM ogistic Regression	Accuracy 47.69736842105263 45.06578947368421 46.05263157894737		Activate Window Go to Settings to activ		
			Model Type leighborsCl <mark>assifier</mark> isision Tree Classifier Naive Bayes SVM ogistic Regression	Accuracy 47.69736842105263 45.06578947368421 46.05263157894737 42.76315789473684 44.07894736842105		Go to Settings to activ		C

Fig 3. Result screenshot 3





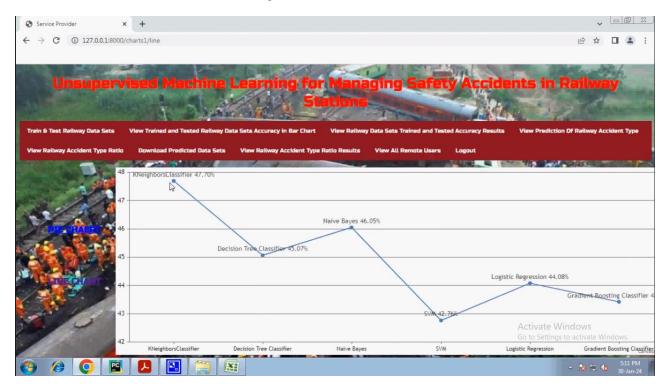


Fig 5. Result screenshot 5

Service Provider ×	+	× [0] [2]
← → C ① 127.0.0.1:8000/c	harts1/pie	🖻 🖈 🔲 🚨 :
Unsupervi	sed Machine Learning for Man Station	aging Safety Accidents in Railway
Train & Test Railway Data Sets	View Trained and Tested Railway Data Sets Accuracy in Bar Chart View Rails	vay Data Sets Trained and Tested Accuracy Results View Prediction Of Railway Accident Type
View Railway Accident Type Ratio	Download Predicted Data Sets View Railway Accident Type Ratio Results	View All Remote Users Logout
	Gradient Boosting Classifier 43.42%	KNeighborsClassifier 47.70% Decision Tree Classifier 45.07%
	Logistic Regression: 44.08%	Naive Bayes 46.05% Activate Windows Go to Settings to activate Windows. Cerves
🚱 🖉 🧿 📓		- 🕅 🐨 🍁 511 PM 30-Jan-24

Fig 6. Result screenshot 6

View Railway Acci	dent Predi	ction Type	Detoils III										_	
RID	Location	Letitude	Longitude	Avg passengers per day		No of trains stopping	No of platforms	No of tracks	Train halting time	Avg trains peed	Average accidents per month	population	P	
10.42.0.42- 54.239.26.255- 33591-443-6	Weifang	36.7167	119.1	9999	68	12	10	1	10	1.6	96	9373000	Pe	
10.42.0.211- 52.84.143.39- 58310-80-6	Fuyang	32.8986	115.8045	952	78	9	2	5	4	1.2	19	8360000	u	
a					N			_						

Fig 7. Result screenshot 7

Vol. 17, Issue.2, 2024

Service Provider × +	~ <u>- @ x</u>
← → C ① 127.0.0.1:8000/View_Predicted_Accident_Type_Ratio/	e 🛧 🛛 😩 :
	and the second
Unsupervised Machine Learning for Managing Safety Ac	cidents in Railway
Stations	
Train 6 Test Railway Data Sets View Trained and Tested Railway Data Sets Accuracy in Bar Chart View Railway Data Sets Trained and Tested Accuracy	y Results View Prediction Of Railway Accident Type
View Rellway Accident Type Ratio Download Predicted Data Sets View Rellway Accident Type Ratio Results View All Remote Users Logout	
View Rallway Accident Prediction Type Ratio Details	1
Railway Accident Type Ratio No Safety Accident 50.0	1.1.
Safety Accident 50.0	
	Activate Windows
	Go to Settings to activate Windows
	- 🔀 🙄 🎝 511 PM

Fig 8. Result screenshot 8

Service Provider ×	+	~		0 2
\leftrightarrow \rightarrow \times (i) 127.0.0.1:8000/View	v_Remote_Users/	6 \$		a :
and the second	eed Machine Learning for Managing Safety Accidents in Ro Stations	K		Type
	Download Predicted Data Sets View Railway Accident Type Ratio Results View All Remote Asers Logout	N -		
USER NAME Gopinath Gop	EMAIL Gender Address Mob No Country State City inath123@gmail.com Male #809,4th Cross,Rajajinagar 9535866270 India Karnataka Bangalore]	
		di.	n	The second
127.0.1:8000/View_Remote_Users/	Activaté Wine Go to Settings to	activate V		AL PM
🔭 🈂 💽 🔛	-)-Jan-24

Fig 9. Result screenshot 9



		OW R DETAILS HERE !!!			
Enter Username	User Name	Enter Password	Password		
Enter EMail Id	Enter Email	Enter Address	Enter Address		
Enter Gender	Select Gender 🗸	Enter Mobile Number	Enter Mobile Number		
Enter Country Name	Enter Country Name	Enter State Name	Enter State Name		
Enter City Name	Enter City Name	1	REGISTER		
	Registe	red Status :			
	Registe	red Status :tu			

Fig 10. Result screenshot 10

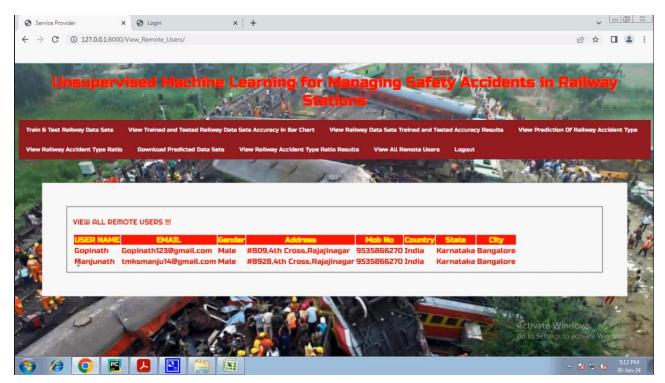


Fig 11. Result screenshot 11



Service Provider	× 🕄 Re	mote User X	+			- <u>- 6 x</u>
\leftrightarrow G (1)	127.0.0.1:8000/ViewYourf	Profile/			07	2 🖬 🏔 E
	Miller	Actide	entș în	Railway Stations		L.R.
	PREDICT RAILWAY	Y ACCIDENT TYPE VIEW Y		DGOUT		
12-1						新学
	YOUR PROFILE DET	rails III				A.
	Username	Manjunath	Email Id	tmksmanju14@gmail.com		-
	Mobile Number	9535866270	Gender	Male		
and the second	Address	#8928,4th Cross,Rajajin	agar Country	India	Activate Windows	1 30
	State	Karnataka	City	Bangalore	Go to Settings to activate	April -
🔭 🙆 (o 🖪 🖊				- 🖪 🖽	5:12 PM 30-Jan-24

Fig 12. Result screenshot 12

Service Provid	er 🗙 🔇 Remote User	×	+		✓ □ ∅ □ □
← → C	① 127.0.0.1:8000/Predict_Accident_Type/				ie 🛧 🗖 🚢 :
			ENTER DATASETS DETAILS HERE III		
	Enter RID		Enter Location		
	Enter Latitude		Enter Longitude		The second
	Enter Avg passengers per day		Enter No of trains passing		200
	Enter No of trains stopping		Enter No of platforms		
/-	Enter No of tracks		Enter Train halting time		
	Enter Avg train speed		Enter Average accidents per mon	th	
	Enter population		Enter Physical Environment		
	Enter DateTime		Enter admin_found		
		Predict			
19				k	
		PREDICTED R	AILWAY ACCIDENT TYPE :: ->	Salety Accident	6.2
9/ 10					
24	111	-			ivate Windows o Settings to activate Windows.
	💽 🗳 🛃 📋				- 😼 🖶 🌆 5:14 PM

Fig 13. Result screenshot 13





Fig 14. Result screenshot 14

Service Provider	Remote User × +	
C 9 C 0 127.00.12	outury view_Predicted_Accident_type_kattor	
Unsupe	rvised Machine Learning for Managing Safety Accidents	in Railway
Train & Test Railway Data Set	View Trained and Tested Railway Data Sets Accuracy in Bar Chart View Railway Data Sets Trained and Tested Accuracy Results View I	Prediction Of Railway Accident Type
View Railway Accident Type R	atio Download Predicted Data Sets View Railway Accident Type Ratio Results View All Remote Lisers Logout	Here and the second second
	View Railway Accident Prediction Type Ratio Details	-
	Railway Accident Type Ratio No Safety Accident 33.3333333333333333333333333333333333	
	Safety Accident 66.65666666666666666	
		vate Windows
7		Settings to activate Windows.
📀 🖉 🚺	S 😕 🔛 🤮 🚳	▲ 10 10 10 10 10 10 10 10 10 10 10 10 10

Fig 15. Result screenshot 15

Service Provider ×	S Remote User	× +			~	
← → C ① 127.0.0.1:8000/cha	arts/line				6 🕯	□ 😩 :
						-
Unsunaryis	sed Machine	Learning for Mar	ading Safet	Accidents in	Dailwa	a h
	and the second s	Station		14		
100	1 200		ST THE REAL	made the internet	7 Parts	ADECH AN
Train & Test Railway Data Sets Vi	'iew Trained and Tested Railway D	lata Sets Accuracy in Bar Chart View Rai	lway Data Sets Trained and Teste	d Accuracy Results View Prediction	n Of Railway Ac	cident Type
View Railway Accident Type Ratio	Download Predicted Data Sets	View Railway Accident Type Ratio Result	ts View All Remote Users	Logout		
	70 -	11	Contraction of the local division of the loc		• 170 Kalon T	TTA
100 M						
Contraction of the second	65					
	60					
CALL DONAL	55			/		
	50					
MARCH NO.	45					
	40	/				
S. M. C.						
	35			Activate V		
	30	No Safety Accident		Go to Setting Safety Acc	to activate W	indows.
		3			- 18 19 4	5:14 PM

Fig 16. Result screenshot 16

Service Provider ×	S Remote User	× +			- <u>- 8 %</u>
← → C ① 127.0.0.1:8000/c	harts/pie				ල් 🕁 🔲 😩 :
Unsupervi	ised Machine I	Learning for Ma	naging Safe	ty Accidents	in Railway
Train & Test Railway Data Sets	View Trained and Tested Railway Da	ta Sets Accuracy in Bar Chart View R	ilway Data Sets Trained and Tes	ited Accuracy Results View F	Prediction Of Railway Accident Type
View Railway Accident Type Ratio	Download Predicted Data Sets	View Railway Accident Type Ratio Resu	lts View All Remote Users	Logout	
		Safety	Lacident: 66.67%	No Safety Accident 33.33	s rate Windows
		Safety Accident 66	67%		Settings to activate Windows.
🚱 🤌 💽 📓					← 💦 💬 🍁 5:14 PM 30-Jan-24

Fig 17. Result screenshot 17

Vol. 17, Issue.2, 2024

Paste	Cut Copy		Arial B I U -	• 10 • 1				Wrap Text		eneral	*	Conditional Formatting *	Format		sert Delete	Format	S	ort & Find	84
*	Format Clipboard	Painter	F		-		Alignment		5	Number	G		styles	Styles *	Cells	- 20	Editir	ilter * Selec	d
	A1		(fx	A87			Saymaca		0.60	ingino ci			sques	1	66113		Contra		
		A		В	С	D	E	F	G	Н	1	J	K	L	M	N	0	P	Q
10	12 0 12 51	220.20	255-33591-443-		50 7407	119.1	9999	68	12	10	7	10	1.6	96	9373000	Dedestrain	02 44 22 0		No. Cofee: Anolder
			.39-58310-80-6		32.8986	115.8045	9999	78	9 45	2 5	5 4	4	1.2	19		Lineside			No Safety Accident Safety Accident
			195-56360-443-			105.5737	7850	80	45	5	4	3	1.8	1	3556000	Road			Safety Accident
3																			
2																			
3																			
												¢							
												-							
																	1	e de	
E.																	tivate V		
	H sheet	. (01)											4			Go Go	to Setting	s to activ	ate Windows.

Fig 18. Result screenshot 18

Overall, the results and discussion underscore the transformative potential of unsupervised machine learning techniques in railway industry safety and other safety-critical domains. By leveraging AI methods such as unsupervised topic modeling, safety administrators can gain a deeper understanding of safety risks and implement targeted interventions to mitigate accidents and enhance safety standards. The study's findings emphasize the importance of utilizing advanced analytical tools to harness the wealth of textual data available in accident reports and other safety-related documents. Through systematic analysis and predictive modeling, railway stakeholders can proactively address safety concerns, reduce the occurrence of accidents, and ensure the continued reliability and safety of railway operations. Moreover, the study highlights the broader implications of AI applications in enhancing safety across various industries, underscoring the potential for future research and innovation in safety management practices.

CONCLUSION

Topic models have an important role in many fields and in such case of safety and risk management in the railway stations for texts mining. In Topic modeling, a topic is a list of words that occur in statistically significant methods. A text can be voice records investigation reports, or reviews risk documents and so on. This research displays various cases for the power of unsupervised machine learning topic modeling in promoting risk management, safety accidents investigation and restructuring accidents recording and documentation on the industry-based level. The description of the root causes accident, the suggested model, it has been showing that the platforms are the hot point in the stations. The outcomes reveal the station's accidents to be occurring owing to four main causes: falls, struck by trains, electric shock. Moreover, the night time and days of the week seems to contact to the risks are significant. With increased safety text mining, knowledge is gained on a wide scale and different periods resulting in greater efficiency RAMS and providing the creation of a holistic perspective for all stakeholders.

Vol. 17, Issue.2, 2024

REFERENCES

1. Chen, Y., & Gao, J. (2016). Unsupervised machine learning: a review. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 6(4), 235-247. https://doi.org/10.1002/widm.1198

2. Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. Journal of Machine Learning Research, 3(Jan), 993–1022. http://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf

3. Dong, J., Wang, X., Zhao, T., & Sun, L. (2020). Application of unsupervised machine learning in mechanical fault diagnosis: A review. IEEE Access, 8, 198819-198833. https://doi.org/10.1109/ACCESS.2020.3033807

4. Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). stm: An R package for structural topic models. Journal of Statistical Software, 91(2), 1-40. https://doi.org/10.18637/jss.v091.i02

5. Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20(1), 37–46. https://doi.org/10.1177/001316446002000104

6. Wang, C., & Blei, D. M. (2011). Collaborative topic modeling for recommending scientific articles. Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 448-456. https://doi.org/10.1145/2020408.2020482

7. Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. Proceedings of the National Academy of Sciences, 101(Suppl 1), 5228–5235. https://doi.org/10.1073/pnas.0307752101

8. Chen, L., Shi, J., & Liu, C. (2014). Applications of unsupervised learning in bridge engineering: a review. Archives of Computational Methods in Engineering, 21(2), 173-192. https://doi.org/10.1007/s11831-013-9096-5

9. Ramage, D., Hall, D., Nallapati, R., & Manning, C. D. (2009). Labeled LDA: A supervised topic model for credit attribution in multi-labeled corpora. Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing, 248–256. https://doi.org/10.3115/1699510.1699548

10. Blei, D. M., & Lafferty, J. D. (2006). Dynamic topic models. Proceedings of the 23rd International Conference on Machine Learning, 113-120. https://doi.org/10.1145/1143844.1143859

11. Mimno, D., Wallach, H. M., Talley, E., Leenders, M., & McCallum, A. (2011). Optimizing semantic coherence in topic models. Proceedings of the Conference on Empirical Methods in Natural Language Processing, 262–272. https://www.aclweb.org/anthology/D11-1020.pdf

12. Buntine, W. (2012). Operations for learning with graphical models. Journal of Machine Learning Research, 13, 1113–1154. http://www.jmlr.org/papers/volume13/buntine12a/buntine12a.pdf

13. Hall, D., Ramage, D., Jurafsky, D., & Manning, C. D. (2008). Studying the history of ideas using topic models. Proceedings of the Conference on Empirical Methods in Natural Language Processing, 363–371. https://doi.org/10.3115/1613715.1613765

14. Blei, D. M., Griffiths, T. L., Jordan, M. I., & Tenenbaum, J. B. (2004). Hierarchical topic models and the nested Chinese restaurant process. Advances in Neural Information Processing Systems, 16, 17–24. https://papers.nips.cc/paper/2004/file/5a6d9264fb2a0db98fb9cf65b0d4a5e7-Paper.pdf

Vol. 17, Issue.2, 2024

15. Zhu, X., & Ghahramani, Z. (2002). Learning from labeled and unlabeled data with label propagation. Technical Report, Carnegie Mellon University, Pittsburgh, PA. https://www.cs.cmu.edu/~zhuxj/pub/CMU-CALD-02-107.pdf

16. Paisley, J., Blei, D. M., & Jordan, M. I. (2012). Variational Bayesian inference with stochastic search. Proceedingsofthe29thInternationalConferenceonMachineLearning,1347–1354.https://dl.acm.org/doi/10.5555/3042573.3042736

17. Wang, X., & McCallum, A. (2006). Topics over time: A non-Markov continuous-time model of topical trends. Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 424–433. https://doi.org/10.1145/1150402.1150469