



National University of Computer and Emerging Sciences, Lahore



Landslide Insight: Exploring Data Patterns and Predictive Modeling

Sophia Razzaq 21L-5607

Idrees Arshad 21L-5632

Burhan Anwar 21L-5676

Aroob Khalid 20L-1313

Instructor: Miss Esha Tur Razia Babar

FAST School of Computing

National University of Computer and Emerging Sciences

Lahore, Pakistan

May 2024

Contents

Introduction	ii
Challenges	ii
Getting to Know the Dataset	iii
Shape of the Data	iii
Column Names and Data Types	iii
Preprocessing	iv
Exploratory Data Analysis and Visualization	v
Global Dataset	v
Pakistan Dataset	x
Machine Learning Algorithms Applied	xiv
SVM	xiv
Bagging	xiv
Gradient Boosting Classifier	xv
AdaBoost Classifier	xvi
Kmeans	xvii
Logistic Regression	xviii
Decision Tree	xviii
Random Forest	xix
LIVE PREDICTION	xx
Conclusion	xxi

Introduction

In this project, the objective was to develop a landslide prediction system using geographical coordinates (longitude and latitude) and the corresponding landslide category. The project began by utilizing the NASA Global Landslide Catalog dataset, which provided a comprehensive collection of landslide occurrences worldwide. Through initial data visualization and analysis, the focus was narrowed down to the region of Pakistan, where the dataset was further explored. Several machine learning algorithms were employed, including DBSCAN, SVM, Bagging, Gradient Boosting Classifier, AdaBoost Classifier, FP Growth, Logistic Regression, Linear Regression, Naive Bayes, Decision Tree, and Random Forest. Each algorithm was evaluated using various metrics to determine its effectiveness in predicting landslide types. The results obtained from these evaluations showcased promising performance, with several algorithms achieving high accuracy percentages. This report presents a detailed account of the methodology, findings, and insights gained from this landslide prediction project, highlighting the potential for utilizing machine learning techniques in improving landslide risk assessment and management.

Challenges

One of the primary challenges was the quality and completeness of the dataset, requiring thorough cleaning and preprocessing. Dealing with missing values, inconsistencies, and outliers was crucial for accurate analysis and modeling. Additionally, imbalanced data distribution posed a challenge, requiring exploration of techniques like oversampling and undersampling for balanced predictions. Selecting relevant features and engineering new ones, considering factors like topography, climate, vegetation, and soil properties, required domain knowledge and careful validation.

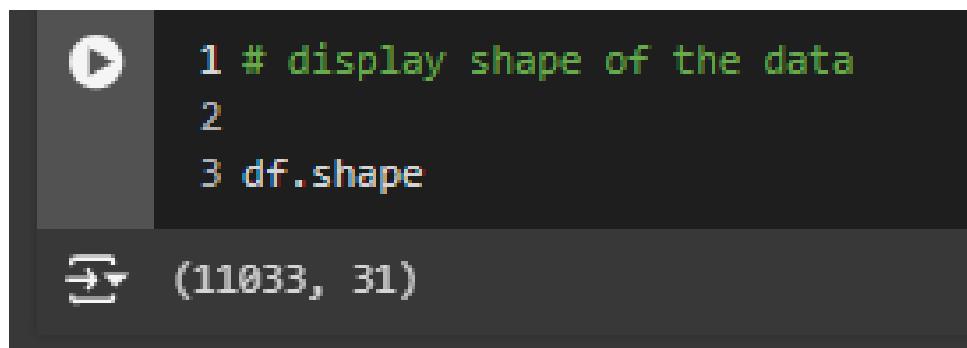
Tawangmangu, central Java																										
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
16	4983	06/25/2013	0:00:0:0	0:0:0:0	0:0:0:0	0:0:0:0	0:0:0:0	0:0:0:0	landslide	downpour	medium	unknown	0	0	0	0	0	gfc	4983	India	IN	Tamil Nid	31364	Kangayam	6.0388	
17	5102	07/14/2013	12:00:0	0:0:0	0:0:0	0:0:0	0:0:0	0:0:0	landslide	unknown	medium	unknown	0	0	0	0	0	gfc	5102	Nepal	NP	Mid West	0	SalyAn	5.9897	
18	5122	07/17/2013	12:00:0	0:0:0	0:0:0	0:0:0	0:0:0	0:0:0	landslide	rain	medium	unknown	0	0	0	0	0	gfc	5122	India	IN	NaGaland	92113	Kohima	19.796	
19	5103	07/14/2013	12:00:0	0:0:0	0:0:0	0:0:0	0:0:0	0:0:0	landslide	unknown	medium	unknown	0	0	0	0	0	gfc	5103	Nepal	NP	Mid West	0	SalyAn	37.559	
20	5123	07/17/2013	12:00:0	0:0:0	0:0:0	0:0:0	0:0:0	0:0:0	landslide	rain	medium	unknown	0	0	0	0	0	gfc	5123	India	IN	NaGaland	92113	Kohima	4.8184	
21	5153	07/21/2013	12:00:0	0:0:0	0:0:0	0:0:0	0:0:0	0:0:0	landslide	rain	medium	unknown	0	0	0	0	0	gfc	5153	Philippines	PH	Central Vis	798634	Cebu City	1.3901	
22	9943	06/13/2017	03:17:0	5:0:0	5:0:0	5:0:0	5:0:0	5:0:0	landslide	Chittagong	25km	continuo	large	urban	33	0	0	0	0	0	0	0	0	0		
23	5597	*****	*****	*****	*****	*****	*****	*****	landslide	rain	medium	unknown	1	1	1	1	1	gfc	5597	Austria	AT	Tyrol	8818	Kitzbühel	0.3382	
24	7056	05/17/2015	12:00:0	0:0:0	0:0:0	0:0:0	0:0:0	0:0:0	landslide	rain	medium	urban	0	0	0	0	0	gfc	7056	Mexico	MX	Baja Calif	1173	La Espera	6.4615	
25	9988	*****	*****	*****	*****	*****	*****	*****	landslide	At least 100m	large	unknown	0	0	0	0	0	0	0	0	0	0	0	0		
26	9989	*****	*****	*****	*****	*****	*****	*****	landslide	shall	transient	large	0	0	0	0	0	0	0	0	0	0	0	0		
27	9987	06/30/2017	04:00:0	0:0:0	0:0:0	0:0:0	0:0:0	0:0:0	landslide	Lebak Kar	50km	continuo	medium	natural	4	2	2	2	2	https://i0.wp.com/www.nepalmountainnews.com/cms/wp-content/uploads/2016/09/Paintedanda-Lek	https://i0.wp.com/www.nepalmountainnews.com/cms/wp-content/uploads/2016/09/Paintedanda-Lek	https://i0.wp.com/www.netralnews.com/foto/2017/06/30/492-hujan_deras_sebabnya_longsor_di_lokal_lo	https://i0.wp.com/www.netralnews.com/foto/2017/06/30/492-hujan_deras_sebabnya_longsor_di_lokal_lo			
28	9984	*****	*****	*****	*****	*****	*****	*****	landslide	Landslide	Wokha Sa	5km	downpour	small	above_rain	0	0	0	0	0	0	0	0	0		
29	10000	*****	*****	*****	*****	*****	*****	*****	landslide	Landslide	Ukhuwa Up	25km	continuo	medium	natural	1	1	1	1	1	Rest of family injured, unsure of number	1	1	1		
30	9940	*****	*****	*****	*****	*****	*****	*****	landslide	Patali	Landslide	Panki	1km	continuo	medium	natural	1	1	1	1	1	http://images.tribunnews.com/gall_content/2017/07/17/75largein03_Monday_2017_171821	http://images.tribunnews.com/gall_content/2017/07/17/75largein03_Monday_2017_171821	http://www.vnews.com/getattachment/76a208e-9010-4eb1-86c5-4981f78545a/Stormolo-rw-vn-07	http://www.vnews.com/getattachment/76a208e-9010-4eb1-86c5-4981f78545a/Stormolo-rw-vn-07	
31	9967	*****	*****	*****	*****	*****	*****	*****	landslide	Heavy rain	exact	debris	0	0	0	0	0	0	0	0	0	0	0	0		
32	9968	*****	*****	*****	*****	*****	*****	*****	landslide	transient	large	0	0	0	0	0	0	0	0	0	0	0	0	0		
33	9941	*****	*****	*****	*****	*****	*****	*****	landslide	Chittagong	Despatch	50km	continuo	small	below_rain	0	0	0	0	0	http://www.dailysun.com/assets/news/_images/2017/07/03/hill-slide.jpg	http://www.dailysun.com/assets/news/_images/2017/07/03/hill-slide.jpg	http://www.dailysun.com/assets/news/_images/2017/07/03/hill-slide.jpg	http://www.dailysun.com/assets/news/_images/2017/07/03/hill-slide.jpg		
34	9969	*****	*****	*****	*****	*****	*****	*****	Rockslide	Tenaga e	Narrows	exact	rock	fall	no_apari	small	natural	1	1	1	1	1	1	1	1	
35	9950	*****	*****	*****	*****	*****	*****	*****	landslide	Three mer	Taplejung	25km	landslide	rain	large	0	4	10	10	10	1	1	1	1		
36	9974	*****	*****	*****	*****	*****	*****	*****	landslide	NH-2	bloc	AH1	Mao	1km	continuo	medium	above_rain	0	0	0	0	0	0	0	0	
37	9687	*****	*****	*****	*****	*****	*****	*****	soldiers	z 2	landslide	1	San Miguel	50km	medium	unknown	unknown	unknown	2	1	1	1	1	1	1	1
38	9963	*****	*****	*****	*****	*****	*****	*****	landslide	transient	large	0	0	0	0	0	0	0	0	0	0	0	0	0		
39	9963	*****	*****	*****	*****	*****	*****	*****	Pasang Lh	landslide	Pasang Lh	exact	landslide	downpour	large	0	0	0	0	0	http://im Date very approximate; article says ""Since June 19""	http://im Date very approximate; article says ""Since June 19""	http://im Date very approximate; article says ""Since June 19""	http://im Date very approximate; article says ""Since June 19""		
40	9955	*****	*****	*****	*****	*****	*****	*****	landslide	A couple	Bojpur	1km	landslide	medium	0	2	2	2	2	Article sparse on details	1	1	1	1		
41	9972	*****	*****	*****	*****	*****	*****	*****	Patali	landslide	Chandrap	10km	landslide	continuo	medium	unknown	1	0	0	0	0	Article sparse on details	1	1	1	
42	9959	*****	*****	*****	*****	*****	*****	*****	Landslide	Fir	famil	Bhagwati	5km	landslide	rain	medium	0	0	0	0	0	http://www.hindustantimes.com/rflimage_size_960x540/HT/p2/2017/06/14/Pictures/bangladesh-wes	http://www.hindustantimes.com/rflimage_size_960x540/HT/p2/2017/06/14/Pictures/bangladesh-wes	http://www.hindustantimes.com/rflimage_size_960x540/HT/p2/2017/06/14/Pictures/bangladesh-wes	http://www.hindustantimes.com/rflimage_size_960x540/HT/p2/2017/06/14/Pictures/bangladesh-wes	
43	9951	*****	*****	*****	*****	*****	*****	*****	Landslide	Article ver	Jorabat	A	5km	landslide	downpour	small	unknown	0	0	0	0	0	0	0	0	
44	9952	*****	*****	*****	*****	*****	*****	*****	Landslide	Article ver	ver	ver	ver	landslide	downpour	small	unknown	0	0	0	0	0	0	0	0	
45	9961	*****	*****	*****	*****	*****	*****	*****	Landslide	Landslide	National	8	25km	landslide	downpour	small	above_rain	0	0	0	0	0	0	0	0	
46	9961	*****	*****	*****	*****	*****	*****	*****	Landslide	Landslide	National	8	25km	landslide	downpour	small	above_rain	0	0	0	0	0	0	0	0	

Figure 1: poor quality of dataset .

Getting to Know the Dataset

To begin the analysis, we first explored the dataset to gain a better understanding of its structure and contents. This section presents the key findings related to the shape of the data and the column names.

Shape of the Data



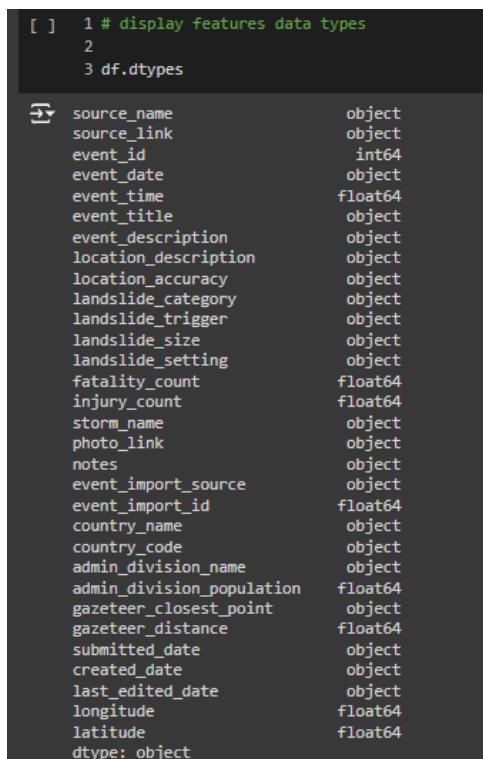
```
1 # display shape of the data
2
3 df.shape
```

2
3 df.shape

2 (11033, 31)

Figure 2: initial shape of dataset .

Column Names and Data Types



```
[ ] 1 # display features data types
2
3 df.dtypes
```

Column Name	Data Type
source_name	object
source_link	object
event_id	int64
event_date	object
event_time	float64
event_title	object
event_description	object
location_description	object
location_accuracy	object
landslide_category	object
landslide_trigger	object
landslide_size	object
landslide_setting	object
fatality_count	float64
injury_count	float64
storm_name	object
photo_link	object
notes	object
event_import_source	object
event_import_id	float64
country_name	object
country_code	object
admin_division_name	object
admin_division_population	float64
gazeteer_closest_point	object
gazeteer_distance	float64
submitted_date	object
created_date	object
last_edited_date	object
longitude	float64
latitude	float64
dtype:	object

Figure 3: initial columns of dataset .

Preprocessing

Removing unwanted Columns

Following columns served no meaning to our prediction and analysis.

```
[ ] 1 # drop unwanted columns
2
3 df.drop(['event_id','event_time','location_description','event_title','event_description','photo_link','notes',
4         'event_import_source','event_import_id','country_code','submitted_date','created_date','last_edited_date'],
5         axis=1,
6         inplace=True)
```

Figure 4: Dropped the unwanted columns

Dealing with date column

```
[ ] 1 # change data type of 'event_date' Column
2
3 df['event_date_cal'] = pd.to_datetime(df['event_date'])

[ ] 1 # split date & time in to separate columns
2
3 df['Date'] = pd.to_datetime(df['event_date_cal']).dt.date
4 df['Time'] = pd.to_datetime(df['event_date_cal']).dt.time

[ ] 1 df.drop(['event_date','event_date_cal'],
2             axis=1,
3             inplace=True)
```

Figure 5: Date Column.

Extracted Date and time from the event column.

Date	Time
2008-08-01	00:00:00
2009-01-02	02:00:00

Figure 6: Updation.

Further advance prepossessing has been done along side when it where it was necessary. At this stage we have completed a generalized prepossessing.

Exploratory Data Analysis and Visualization

GLOBAL DATASET

At first we will see trends around the globe and make interpretations accordingly.

Event Reported Source

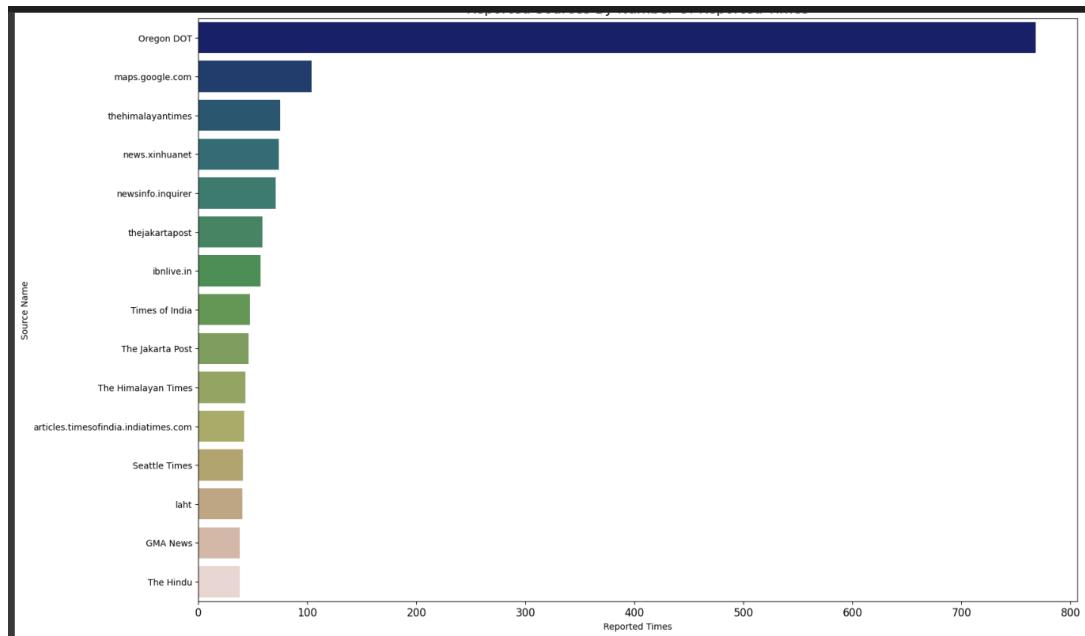


Figure 7: Number of reports made by each source

Origon DOT Have Reported Huge Number Of Events During This Time Period (1988 - 2017)

Geospatial Visualization Of Globally Events



Figure 8: Open Street Map Style

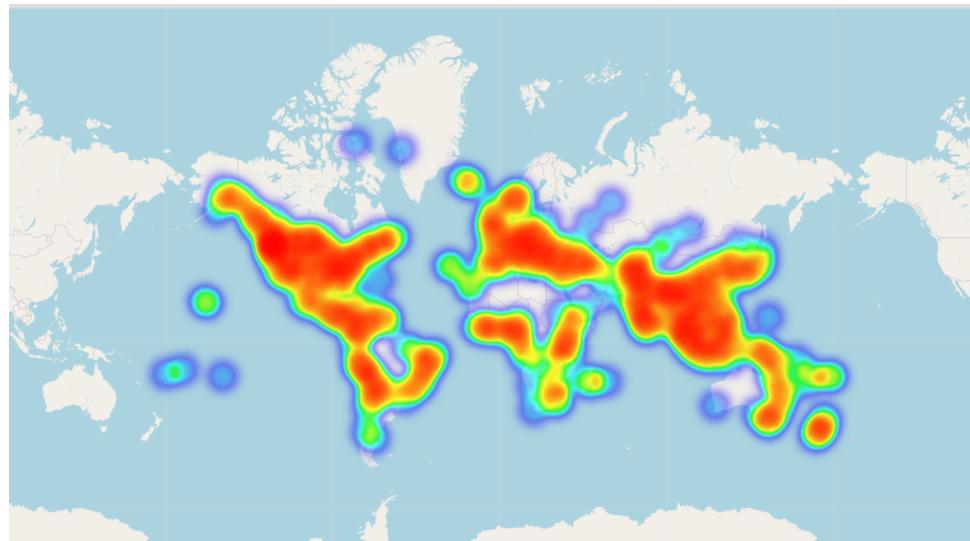


Figure 9: Heatmap.

According to the above maps we can determine that lot of land slide events happen in Indian Ocean
North America South America

Events by year and month

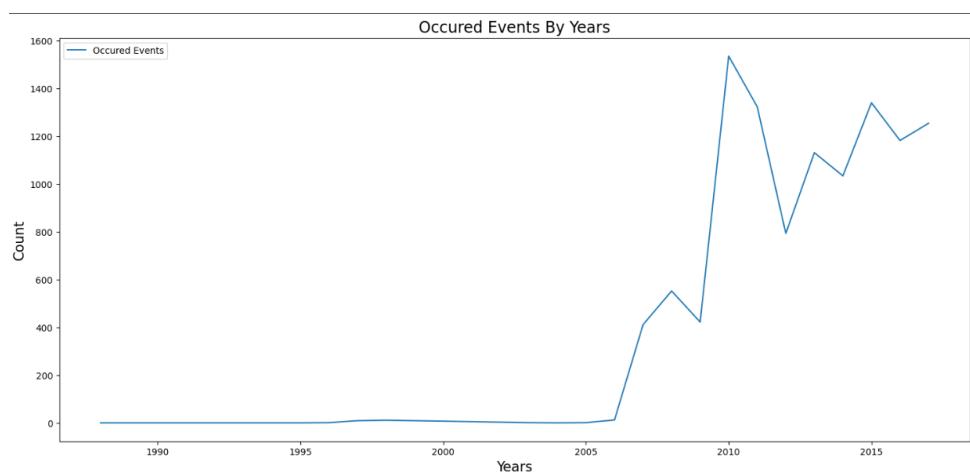


Figure 10: Events by year.

We Can See There Are Most Event Occured During 2010.

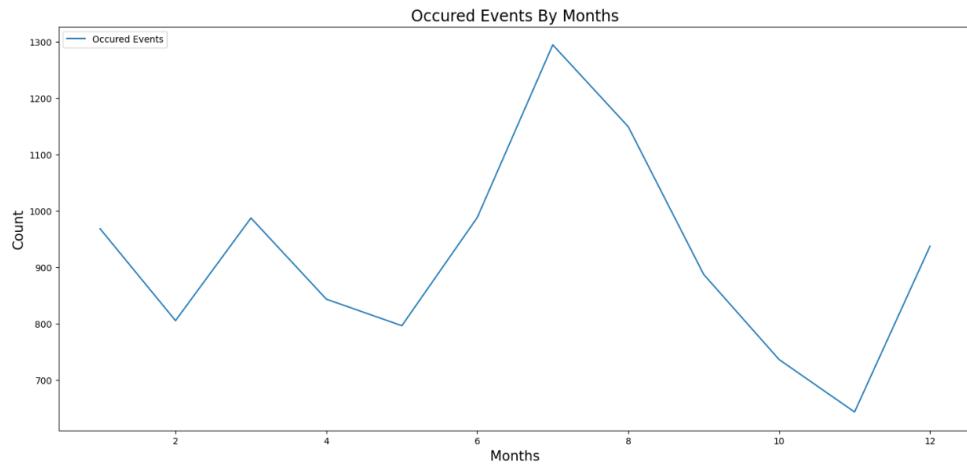


Figure 11: Events by month.

We Can See There Are Most Event Occured During 3 Quater Of The year.

Events in 2010

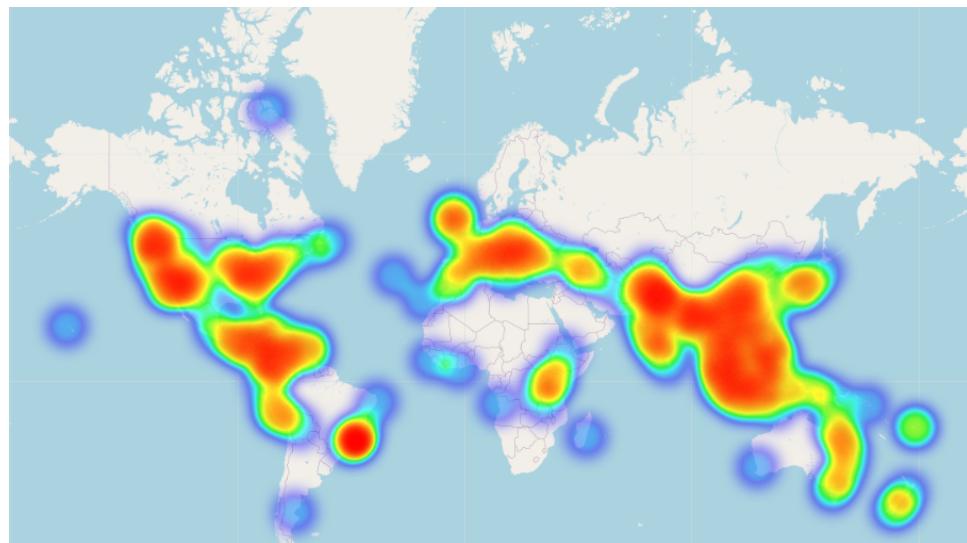


Figure 12: Heatmap of events occurred in 2010

In 2010 More Events Occured In Indian Ocean

Landslide categories

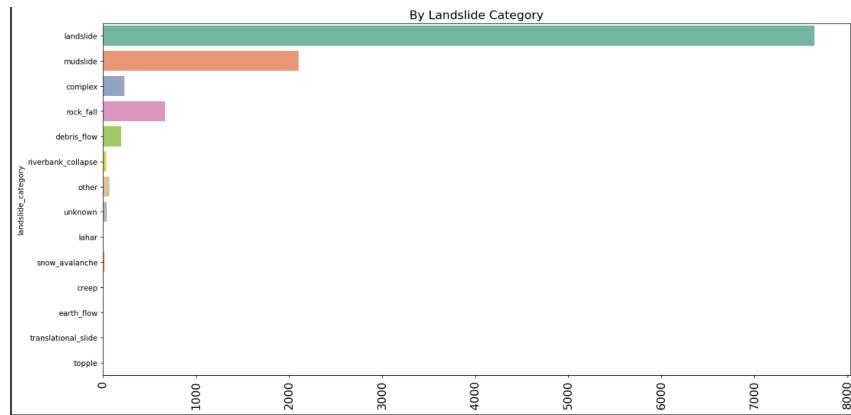


Figure 13: Landslide categories

We can see mostly the news reports were of Landslide.

Event Setting Types

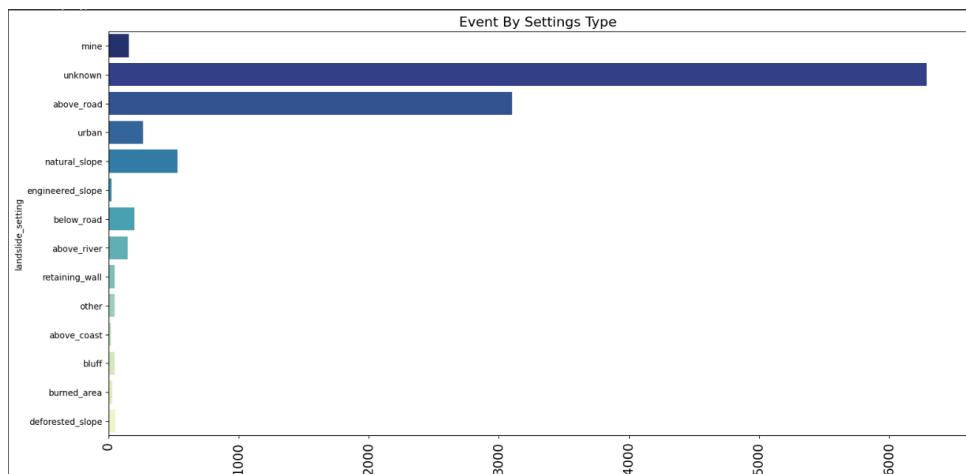


Figure 14: Event Setting Type.

According to above map we can determine lots of events happen in above Roads setting. so roads contructions cause to these events most. if we can construct roads with more safety and pre analysis we can reduce these events happening

Event Trigger Type

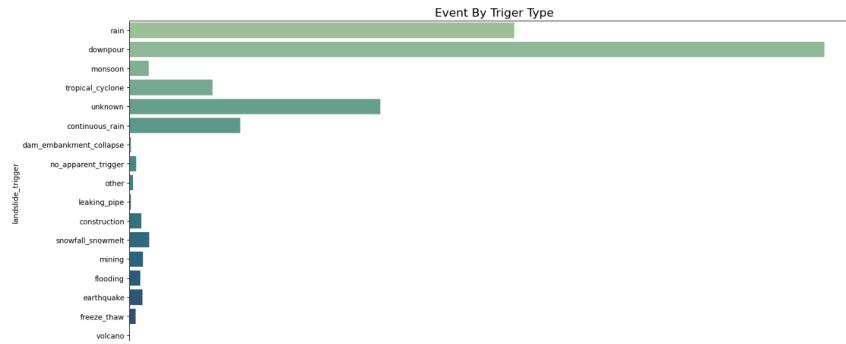


Figure 15: Event Trigger Type.

We can see downpour and raining has triggered landslides the most.

Top 10 countries with most events



Figure 16: Top 10.

Insights

The comprehensive analysis of the global dataset spanning from 1988 to 2017 sheds light on the patterns and trends of landslide events worldwide. Notably, the data highlights the significant reporting role of the origin DOT. Geospatial visualizations pinpoint hotspots primarily in the Indian Ocean, North America, and South America. The peak in landslide occurrences during 2010, particularly in the third quarter, underscores a crucial period for such events. Furthermore, the dominance of landslides, often triggered by downpour and rainfall, underscores the need for proactive measures, especially in road construction, to mitigate future incidents. While US and India remain the highest landslide reported countries by far.

PAKISTAN DATASET

Extracting Pakistan Data



```
[ ] 1 pakistan_df = df[(df['longitude'] < 77.8231) & (df['latitude'] < 37.0841) & (df['latitude'] > 23.6345) & (df['longitude'] > 60.6720)]
```

Figure 17: Pakistan coordinates.

After extracting Pakistan data we found out that there Are 596 Events Occurred In Pakistan During 1988 To 2017.

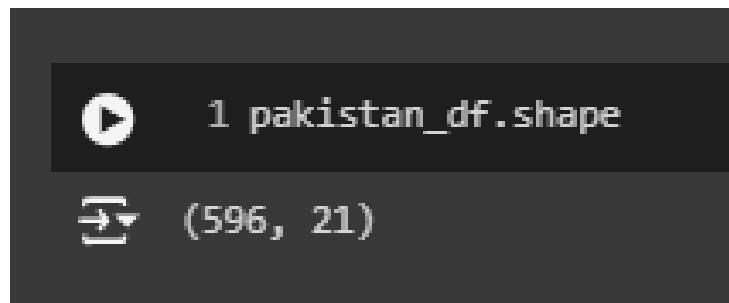


Figure 18: Number of events.

Occured Events In Pakistan Geospatial Visualization

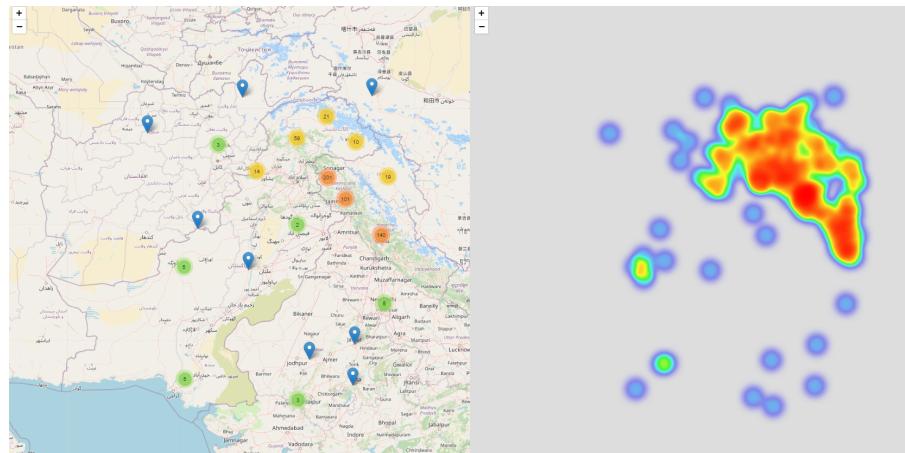


Figure 19: Geospatial Visualization

According To The Above Map, We can See That In Pakistan Most Of The Events Mountain Side Areas. Such As Gilgit Baltistan, KPK and Kashmir.

Time Series Of Occurred Events By Years In Pakistan

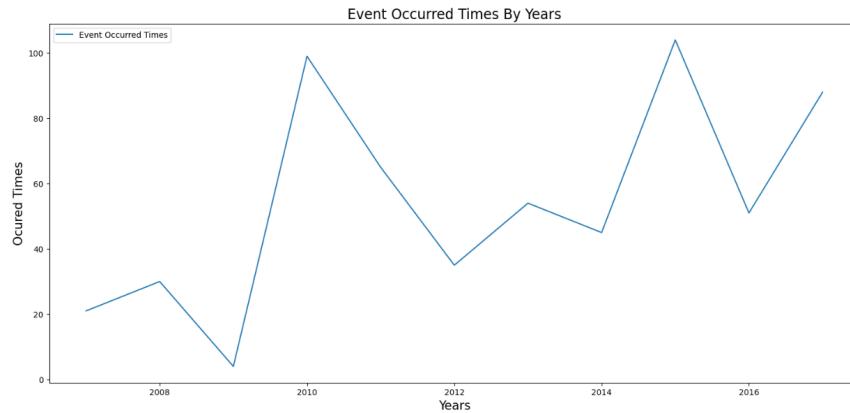


Figure 20: By years

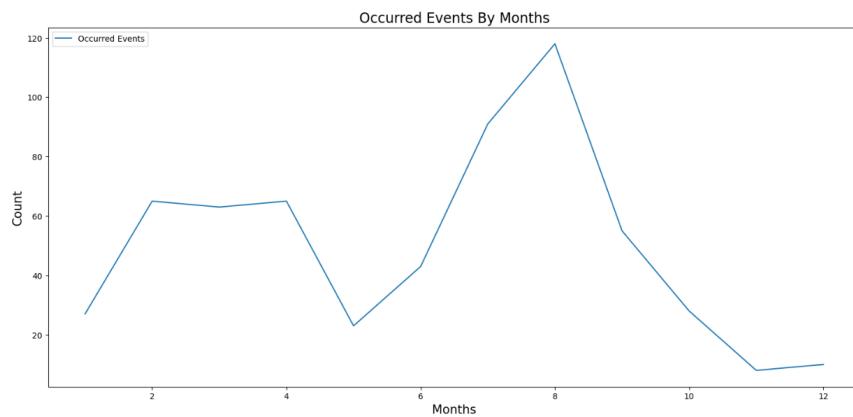


Figure 21: By month

We can see that most events occurred during 2010 and 2015 mostly during august.

Time Series Of Fatalities and Injuries Comparing To Events Occurred Times In Pakistan

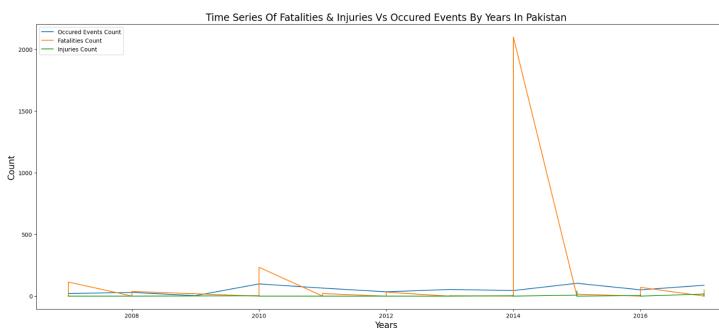


Figure 22: Fatality and Injury count through years

We can observe that the highest fatality count reported were during 2014-2015 which was previously marked as highest events reported year, with over 2k fatality count.

Biggest Landslide Disaster In Pakistan

In 2014, Pakistan experienced one of the biggest landslide disasters in its history. The details of this event can be found in the *Annual Flood Report 2014* provided by the Federal Flood Commission (FFC) of Pakistan. The report can be accessed at the following URL: <https://mowr.gov.pk/SiteImage/Misc/files/2014%20Annual%20Flood%20Report%20of%20FFC.pdf>.

```

Details About Occurred Event With Highest Fatalities
[ ] 1 high_ft = pd.DataFrame(pakistan_df['fatality_count'] == pakistan_df['fatality_count'].max())
2 high_ft
+ Code + Text
source_link location_accuracy landslide_category landslide_trigger landslide_size landslide_setting fatality_count injury_count storm_name ... admin_division
/www.disasternews.net/news/article.php? a... 1km landslide continuous_rain very_large natural_slope 2100.0 0.0 NaN ... Badakshan
f columns (21) exceeds max_columns (20) limiting to #first (20) columns.

```

Figure 23: Event Details

Distribution Of Categories In Occurred Events In Pakistan

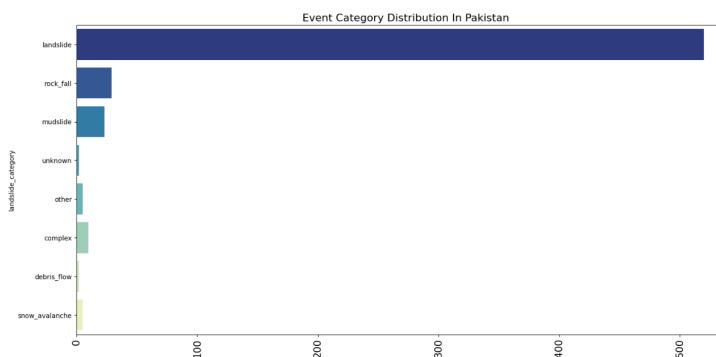


Figure 24: Categories

We can see that majorly Landslides have occurred in Pakistan following rockfall.

Distribution Of Trigger Type In Occurred Events In Pakistan

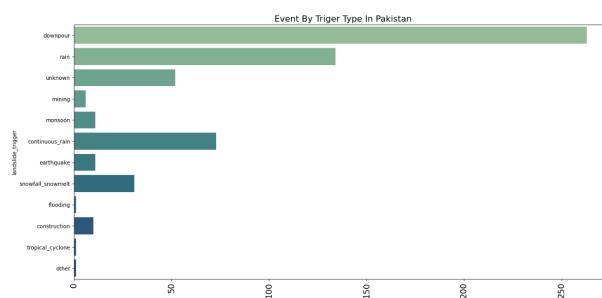


Figure 25: Trigger Types

Downpour and rain has triggered landslides more often.

Distribution Of Event Size In Occurred Events In Pakistan

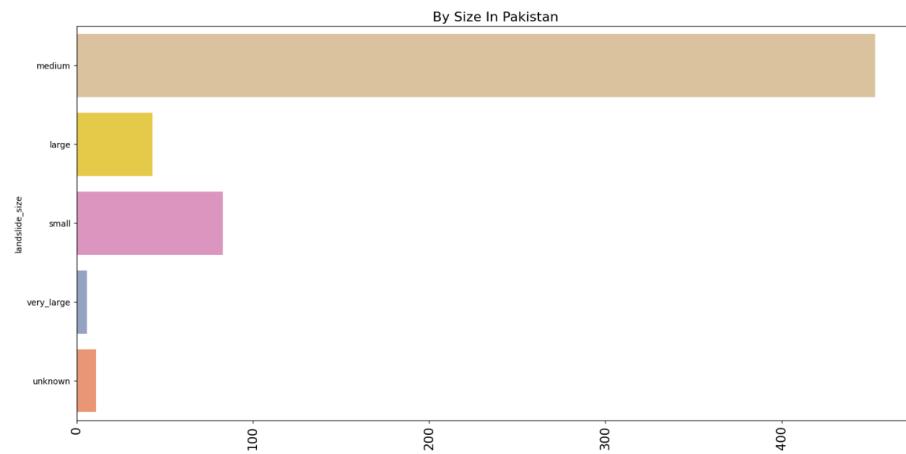


Figure 26: Landslide Size.

Pakistan had a lot of medium sized landslides rather than large scale or small.

Event Reported Sources About In Pakistan

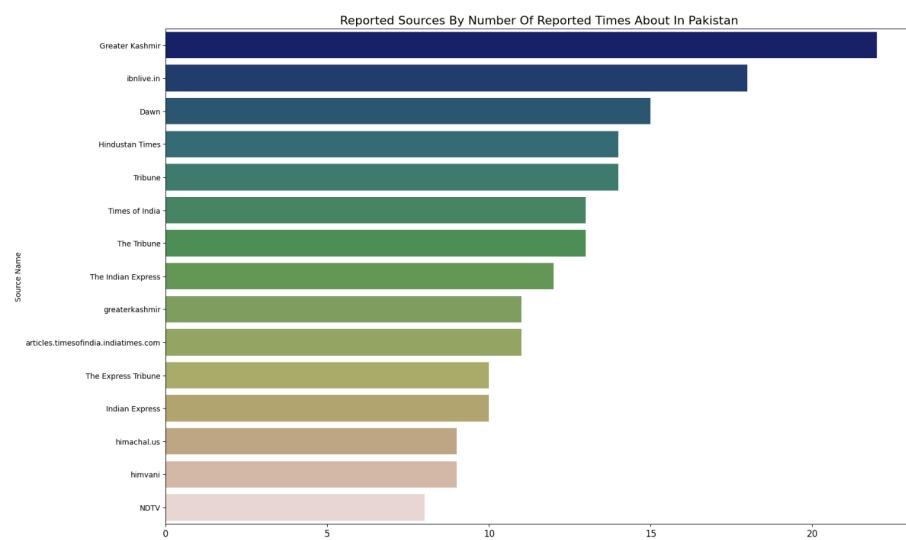


Figure 27: Event Reported Sources About In Pakistan

Greater Kashmir Mirror Has Did Good Job When Reporting Events.

Insights gained

The analysis of Pakistan's landslide data from 1988 to 2017 reveals concentrated occurrences in mountainous regions like Gilgit Baltistan, KPK, and Kashmir. Temporally, peaks in events during 2010 and 2015, particularly in August, are notable. Correlations between event frequencies and consequences, such as fatalities, are evident. Effective communication channels, like Greater Kashmir Mirror, play a vital role. These insights inform targeted risk management strategies and proactive disaster preparedness.

efforts in Pakistan.

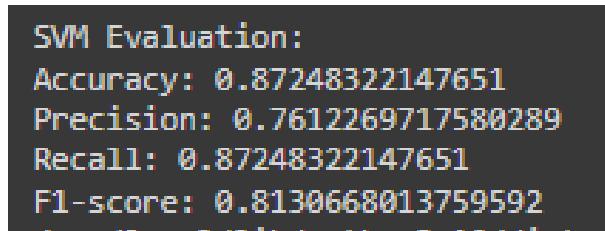
Machine Learning Algorithms Applied

SVM

SVM is effective for classification tasks when there is a clear margin of separation between classes. SVMs can handle both linear and non-linear relationships between features, making them suitable for predicting landslide categories.

```
svm_model = SVC()  
svm_model.fit(X_svm_encoded, y_encoded)  
y_svm_pred = svm_model.predict(X_svm_encoded)
```

Evaluation Metrics



```
SVM Evaluation:  
Accuracy: 0.87248322147651  
Precision: 0.7612269717580289  
Recall: 0.87248322147651  
F1-score: 0.8130668013759592
```

Figure 28: Svm evaluation.

Bagging

Bagging (Bootstrap Aggregating) is an ensemble method that builds multiple base models (in this case, decision trees) on different subsets of the dataset and then combines their predictions. The features chosen for the Bagging Classifier include geographical coordinates ('longitude' and 'latitude'), 'year', and 'month', which are likely to have a significant impact on the occurrence of landslides. These features are chosen because they provide information about the location and time of landslide events, which can be crucial for predictive modeling.

Evaluation Metrics

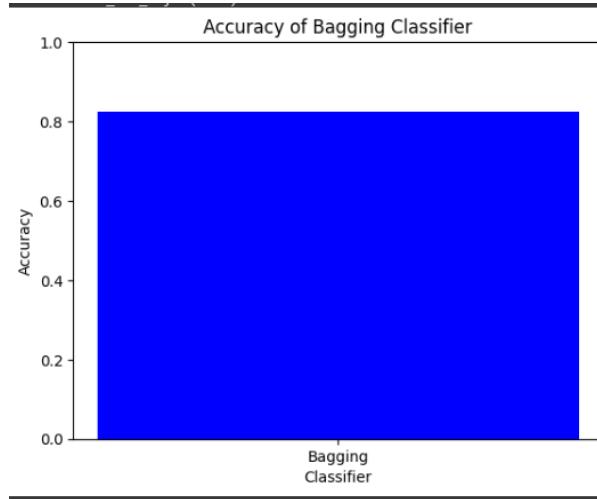


Figure 29: Bagging Results.

Gradient Boosting Classifier

Gradient Boosting is a powerful ensemble method that builds decision trees sequentially, focusing on minimizing the errors of the previous trees. The selected features cover various aspects related to landslide events, such as the triggering factors landslide trigger, the size of the landslide landslide size, the setting landslide setting, information about storms storm name, and the population of administrative divisions admin division population. These features are chosen to capture different dimensions of landslide occurrences and their potential causes.

Evaluation Metrics

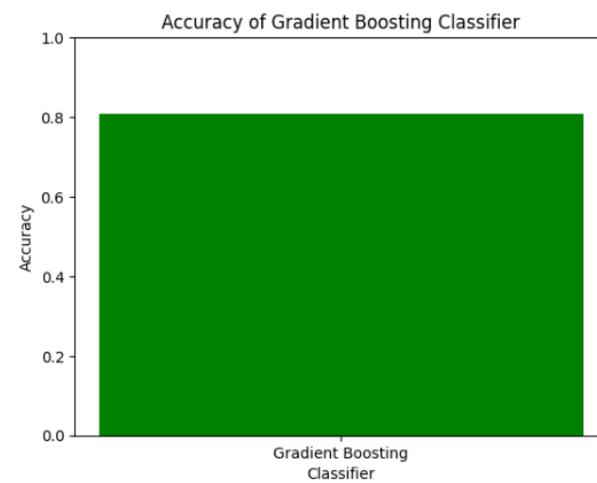


Figure 30: GB Results.

AdaBoost Classifier

Features Used: longitude, latitude, landslide size, admin division population

Reasoning: AdaBoost is an adaptive boosting algorithm that sequentially fits weak learners to the dataset, with a focus on instances that were misclassified by previous learners.

```
# Define and train the AdaBoost classifier
adaboost_classifier = Pipeline(steps=[

    ('preprocessor', preprocessor),
    ('classifier',
     AdaBoostClassifier(base_estimator=DecisionTreeClassifier(),
                        n_estimators=10, random_state=42))

])
```

The selected features include geographical coordinates ('longitude' and 'latitude'), 'landslide size', and 'admin division population'. These features are chosen to provide information about the location and magnitude of landslide events, as well as the population density of administrative divisions, which could influence the occurrence and impact of landslides.

Evaluation Metrics

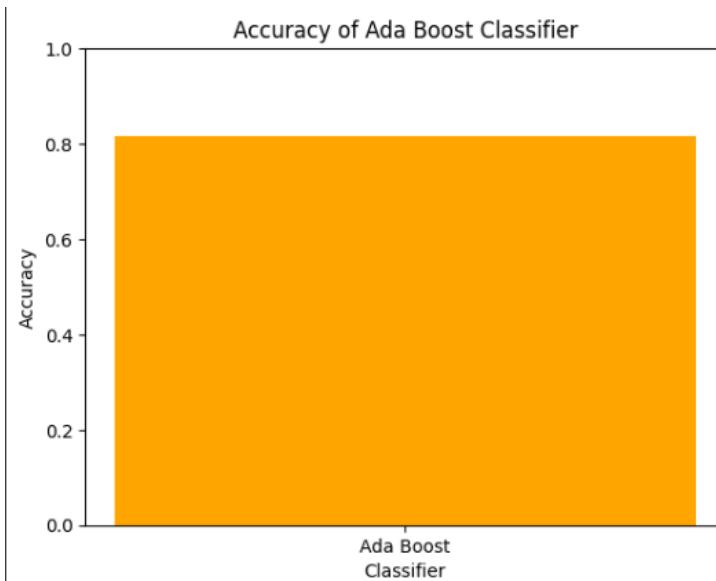


Figure 31: AdaBoosting Results

Kmeans

We used Kmeans to identify natural groupings or clusters of landslides based on their geographical coordinates or other features. K-means clustering helps in understanding spatial patterns in landslide occurrences.

We used Elbow method to determine number of cluster.

```
# Determine the number of clusters using the elbow method
wcss = []
max_clusters = 10 # Maximum number of clusters to consider
for i in range(1, max_clusters + 1):
    kmeans = KMeans(n_clusters=i, init='k-means++', random_state=42)
    kmeans.fit(X)
    wcss.append(kmeans.inertia_)
```

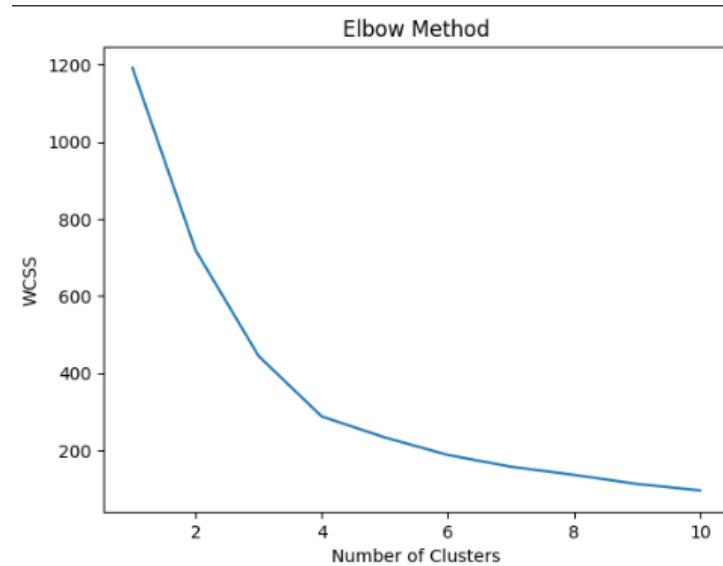


Figure 32: WCSS Plotting

Kmeans Graph

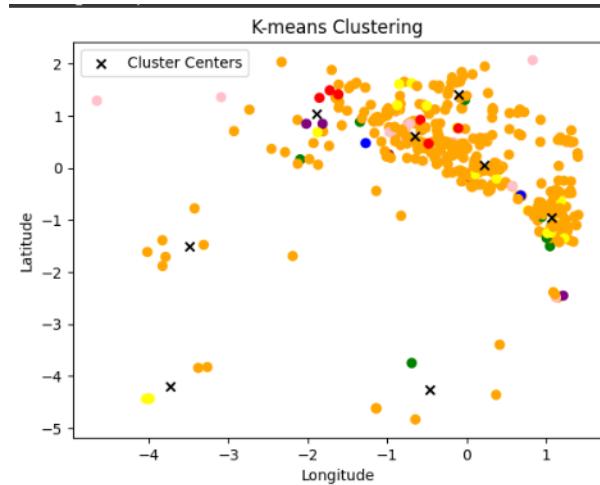


Figure 33: Kmeans Clusters

We can see that each category is plotted to a cluster wrt to longitude and latitude.

Logistic Regression

```
logistic_regression_model = LogisticRegression(C=4, penalty='l2')
logistic_regression_model.fit(X_train, y_train)
```

We applied L2 regularization to the logistic regression model by setting the penalty parameter to 'l2' and the regularization strength parameter (C) to 4. Because previously we noticed that without regularization our model was not performing well.

Evaluation Metrics

```
Accuracy: 0.8166666666666667
Precision: 0.6851694915254237
Recall: 0.8166666666666667
F1-score: 0.7451612903225807
```

Figure 34: Evaluation Matrices.

Decision Tree

We have used to predict the category of landslides based on various features like trigger, size, and setting. Decision trees provide interpretable rules for classification tasks.

```
# Decision Tree with regularization
decision_tree_model = DecisionTreeClassifier(max_depth=8)    ## changing depth
```

```
decision_tree_model.fit(X_dt_rf_encoded, y_encoded)
```

We have used hyper-parameters and change the decision tree depth manually to get better results.

Evaluation Metrics

```
Decision Tree Evaluation:  
Accuracy: 0.912751677852349  
Precision: 0.9109657062530367  
Recall: 0.912751677852349  
F1-score: 0.9096126324995804
```

Figure 35: DT Results.

Random Forest

We have used to predict the category of landslides based on various features like trigger, size, and setting. Decision trees provide interpretable rules for classification tasks.

```
# Random Forest with regularization  
random_forest_model = RandomForestClassifier(max_depth=7)  
random_forest_model.fit(X_dt_rf_encoded, y_encoded)
```

We have used hyper-parameters and change the random forest depth manually to get better results.

Evaluation Metrics

```
Random Forest Evaluation:  
Accuracy: 0.8657718120805369  
Precision: 0.8815095545626126  
Recall: 0.8657718120805369  
F1-score: 0.8438253812091264
```

Figure 36: Random Forest Results.

BEST MODEL

So by using all the models we concluded that decision tree has better results than the other models so we will use this for our predictions.

LIVE PREDICTION

Lets test our decision tree on unseen data

```

51
52
53 longitude = 74.8500
54 latitude = 35.3191
55 year = 2024
56 Month = 3
57 day = 4
58 predicted_category = predict_landslide_category(longitude, latitude, year, Month, day)
59 print("Predicted Landslide Category:", predicted_category)

→ Predicted Landslide Category: landslide
/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning: `should_r
and should_run_async(code)
```

Figure 37: Input Values

Lets test out result through a quick google search and here is what we found out:

One killed as snow, landslide bring life to standstill in GB

Jamil Nagri | Published March 5, 2024



People walk along a snow laden street in Kalam on March 4, 2024. — AFP

Figure 38: Dawn News Article.

So our model predicted an in real life event. Showing how reliable it is!

You may check out the article: [Click here](#)

Conclusion

In conclusion, among the various machine learning algorithms applied, the decision tree stood out as the most effective for predicting landslide categories, achieving an impressive accuracy rate of 91. This high accuracy demonstrates the robustness and reliability of the decision tree model in classifying landslide events based on factors such as trigger type, size, and setting. With its interpretable rules and strong predictive power, the decision tree algorithm emerges as a valuable tool for landslide risk assessment and management.