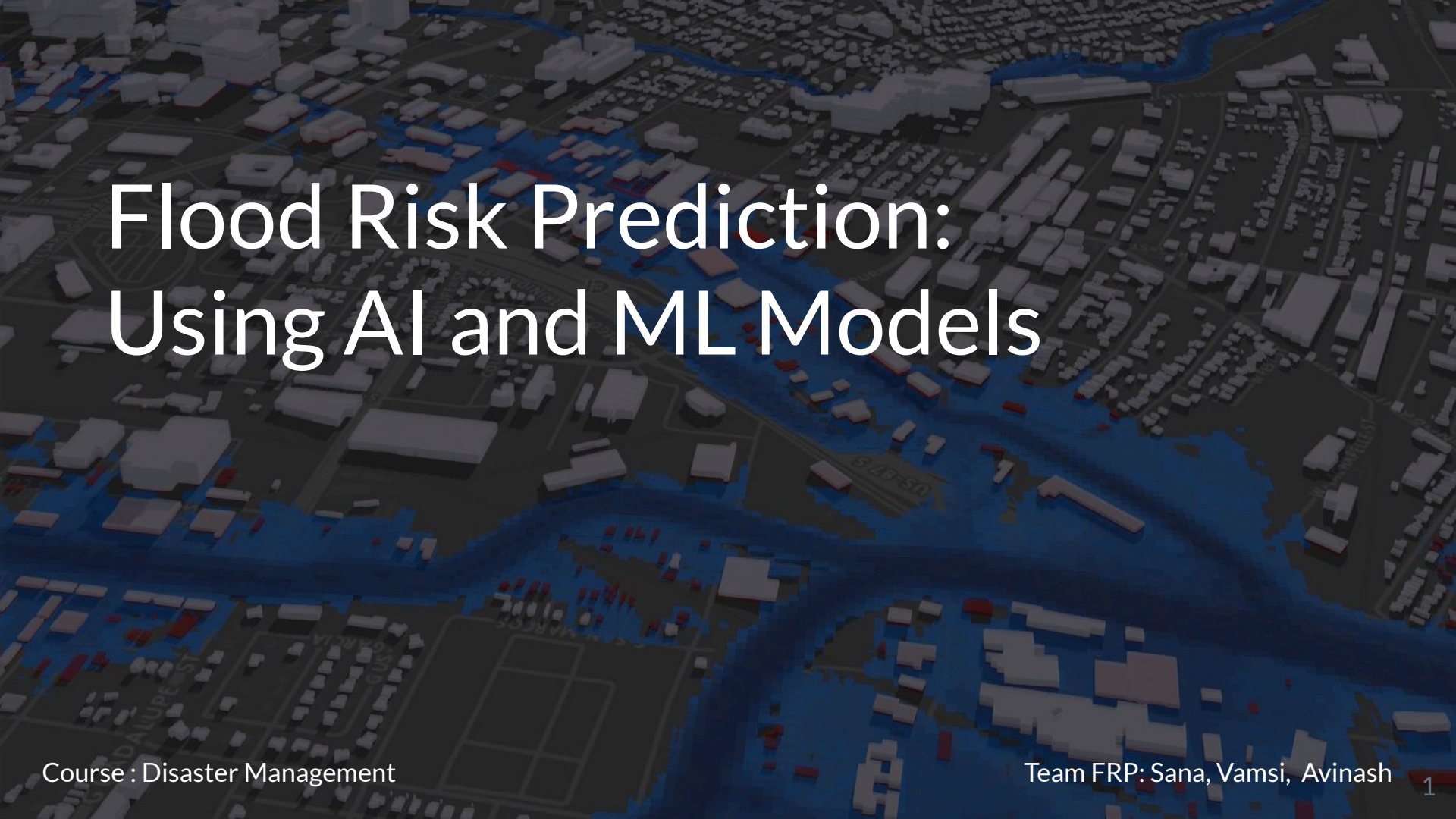


Flood Risk Prediction: Using AI and ML Models

Course : Disaster Management

Team FRP: Sana, Vamsi, Avinash

Introduction

- India is the country in the world that experiences the most catastrophic flooding each year.

- Several essential factors including *floor runoff*, *relative altitude*, and an *insufficient route* for the water to escape are to blame for the increase in water logging.

https://tse4.mm.bing.net/th?id=OIP.Zj3b5SAxx5jj25FGeStAKQHaEI&pid=Api&P=0

Introduction

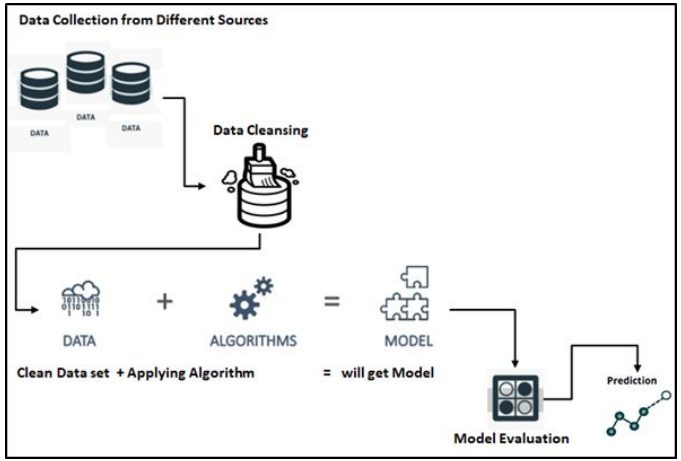
- Forecasting floods is so crucial. In India, flooding is a severe disaster that harms the entire biosphere severely.

- To forecast and lower the risk of flooding, it is essential to design a method for predicting flooding.

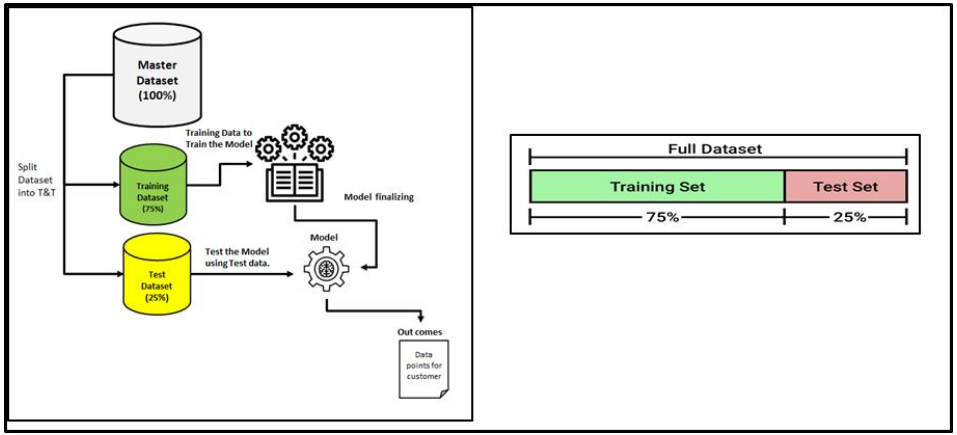
Objective

Use AI/ML algorithms to predict the risk of flooding in a particular area based on historical data

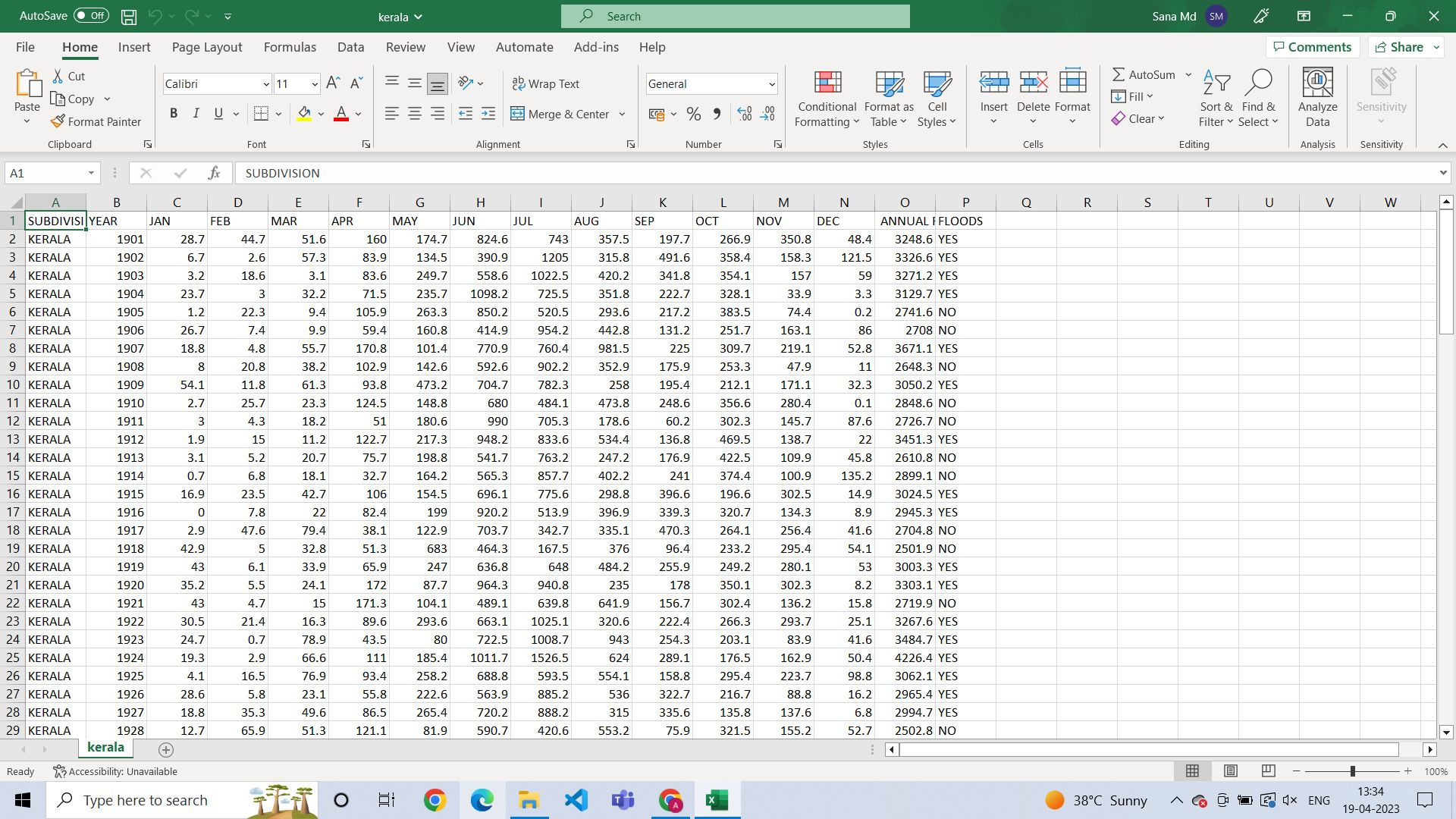
How ML works



https://editor.analyticsvidhya.com/uploads/99744ML%20process.png



https://editor.analyticsvidhya.com/uploads/14391TT.png



AutoSave Off | kerala | Search | Sana Md SM

File Home Insert Page Layout Formulas Data Review View Automate Add-ins Help | Comments | Share

Paste | Cut | Copy | Format Painter | Clipboard

Calibri | 11 | Font

Wrap Text | Merge & Center | Alignment

General | Number

Conditional Formatting | Format as Table | Cell Styles | Styles

Insert | Delete | Format | Cells

AutoSum | Fill | Clear | Sort & Filter | Find & Select | Editing

Analyze Data | Analysis

Sensitivity | Sensitivity

A1 | SUBDIVISION

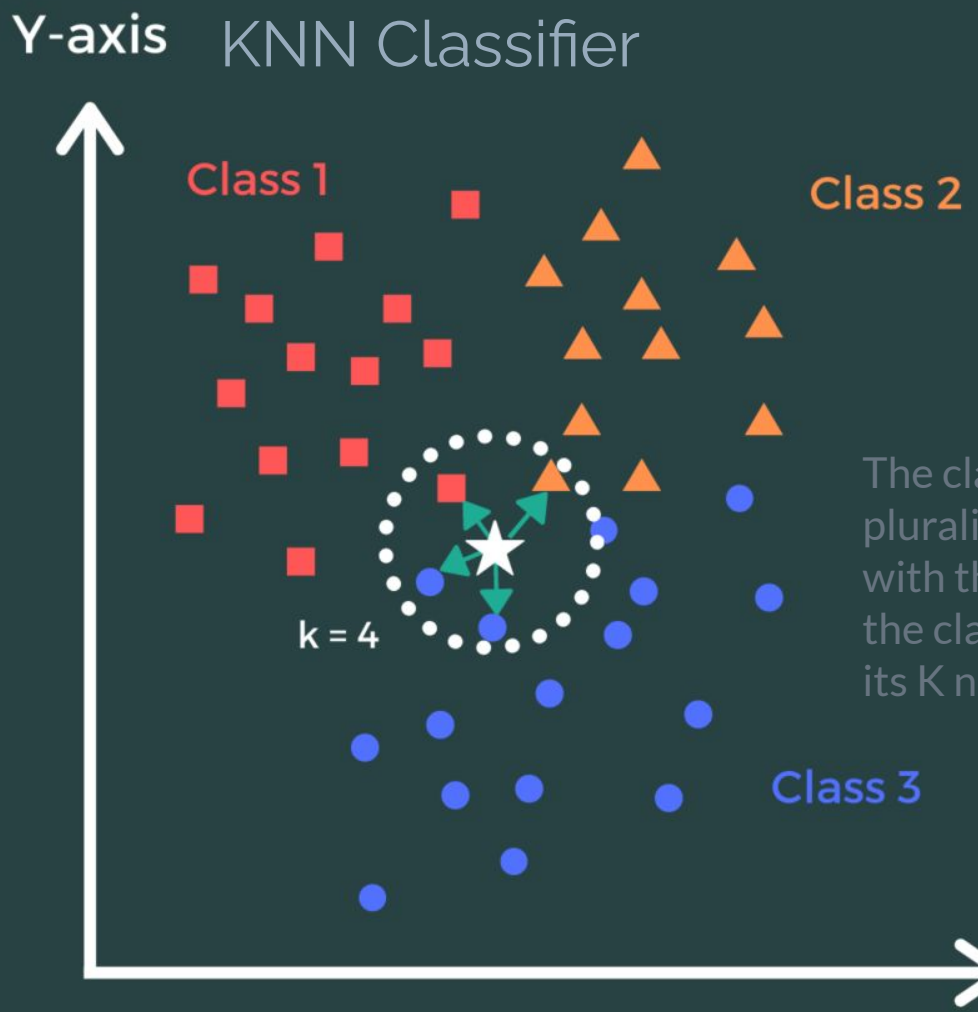
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	SUBDIVISI	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL I	FLOODS
2	KERALA	1901	28.7	44.7	51.6	160	174.7	824.6	743	357.5	197.7	266.9	350.8	48.4	3248.6	YES
3	KERALA	1902	6.7	2.6	57.3	83.9	134.5	390.9	1205	315.8	491.6	358.4	158.3	121.5	3326.6	YES
4	KERALA	1903	3.2	18.6	3.1	83.6	249.7	558.6	1022.5	420.2	341.8	354.1	157	59	3271.2	YES
5	KERALA	1904	23.7	3	32.2	71.5	235.7	1098.2	725.5	351.8	222.7	328.1	33.9	3.3	3129.7	YES
6	KERALA	1905	1.2	22.3	9.4	105.9	263.3	850.2	520.5	293.6	217.2	383.5	74.4	0.2	2741.6	NO
7	KERALA	1906	26.7	7.4	9.9	59.4	160.8	414.9	954.2	442.8	131.2	251.7	163.1	86	2708	NO
8	KERALA	1907	18.8	4.8	55.7	170.8	101.4	770.9	760.4	981.5	225	309.7	219.1	52.8	3671.1	YES
9	KERALA	1908	8	20.8	38.2	102.9	142.6	592.6	902.2	352.9	175.9	253.3	47.9	11	2648.3	NO
10	KERALA	1909	54.1	11.8	61.3	93.8	473.2	704.7	782.3	258	195.4	212.1	171.1	32.3	3050.2	YES
11	KERALA	1910	2.7	25.7	23.3	124.5	148.8	680	484.1	473.8	248.6	356.6	280.4	0.1	2848.6	NO
12	KERALA	1911	3	4.3	18.2	51	180.6	990	705.3	178.6	60.2	302.3	145.7	87.6	2726.7	NO
13	KERALA	1912	1.9	15	11.2	122.7	217.3	948.2	833.6	534.4	136.8	469.5	138.7	22	3451.3	YES
14	KERALA	1913	3.1	5.2	20.7	75.7	198.8	541.7	763.2	247.2	176.9	422.5	109.9	45.8	2610.8	NO
15	KERALA	1914	0.7	6.8	18.1	32.7	164.2	565.3	857.7	402.2	241	374.4	100.9	135.2	2899.1	NO
16	KERALA	1915	16.9	23.5	42.7	106	154.5	696.1	775.6	298.8	396.6	196.6	302.5	14.9	3024.5	YES
17	KERALA	1916	0	7.8	22	82.4	199	920.2	513.9	396.9	339.3	320.7	134.3	8.9	2945.3	YES
18	KERALA	1917	2.9	47.6	79.4	38.1	122.9	703.7	342.7	335.1	470.3	264.1	256.4	41.6	2704.8	NO
19	KERALA	1918	42.9	5	32.8	51.3	683	464.3	167.5	376	96.4	233.2	295.4	54.1	2501.9	NO
20	KERALA	1919	43	6.1	33.9	65.9	247	636.8	648	484.2	255.9	249.2	280.1	53	3003.3	YES
21	KERALA	1920	35.2	5.5	24.1	172	87.7	964.3	940.8	235	178	350.1	302.3	8.2	3303.1	YES
22	KERALA	1921	43	4.7	15	171.3	104.1	489.1	639.8	641.9	156.7	302.4	136.2	15.8	2719.9	NO
23	KERALA	1922	30.5	21.4	16.3	89.6	293.6	663.1	1025.1	320.6	222.4	266.3	293.7	25.1	3267.6	YES
24	KERALA	1923	24.7	0.7	78.9	43.5	80	722.5	1008.7	943	254.3	203.1	83.9	41.6	3484.7	YES
25	KERALA	1924	19.3	2.9	66.6	111	185.4	1011.7	1526.5	624	289.1	176.5	162.9	50.4	4226.4	YES
26	KERALA	1925	4.1	16.5	76.9	93.4	258.2	688.8	593.5	554.1	158.8	295.4	223.7	98.8	3062.1	YES
27	KERALA	1926	28.6	5.8	23.1	55.8	222.6	563.9	885.2	536	322.7	216.7	88.8	16.2	2965.4	YES
28	KERALA	1927	18.8	35.3	49.6	86.5	265.4	720.2	888.2	315	335.6	135.8	137.6	6.8	2994.7	YES
29	KERALA	1928	12.7	65.9	51.3	121.1	81.9	590.7	420.6	553.2	75.9	321.5	155.2	52.7	2502.8	NO

kerala

Ready | Accessibility: Unavailable | 100%

Type here to search | 38°C Sunny | ENG | 13:34 | 19-04-2023

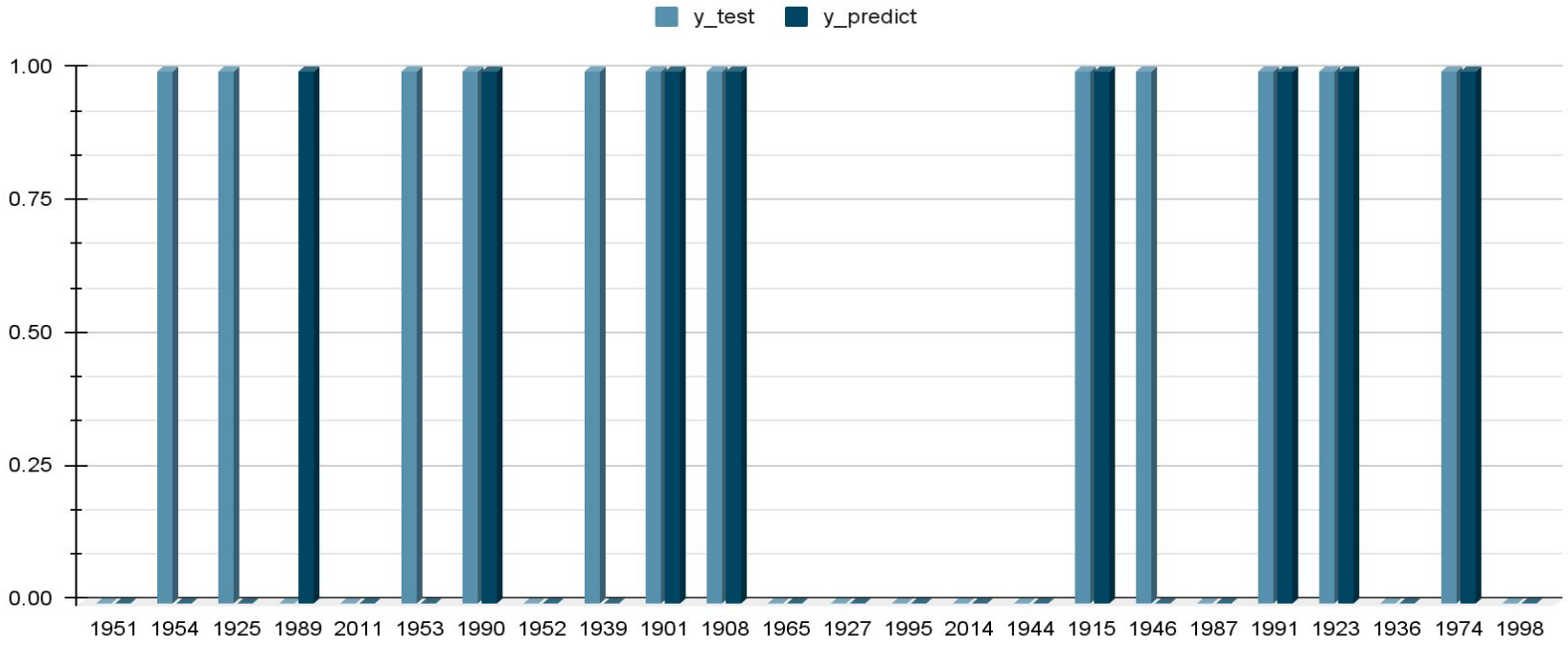
KNN Classifier



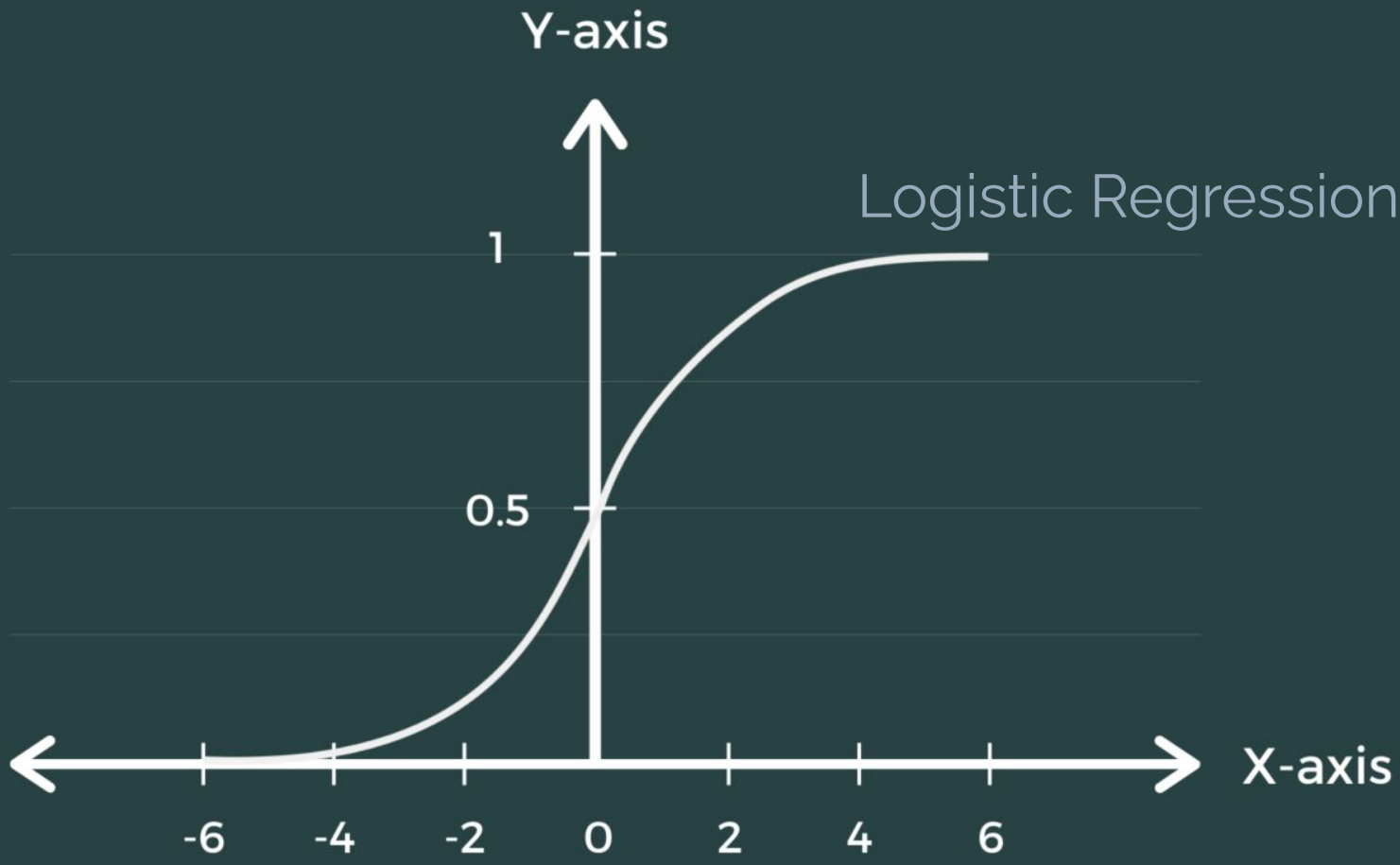
The classification is done by a plurality vote of its neighbours, with the object being assigned to the class most common among its K nearest neighbours.

https://www.theclickreader.com/wp-content/uploads/2020/07/knn-board-1536x864.png

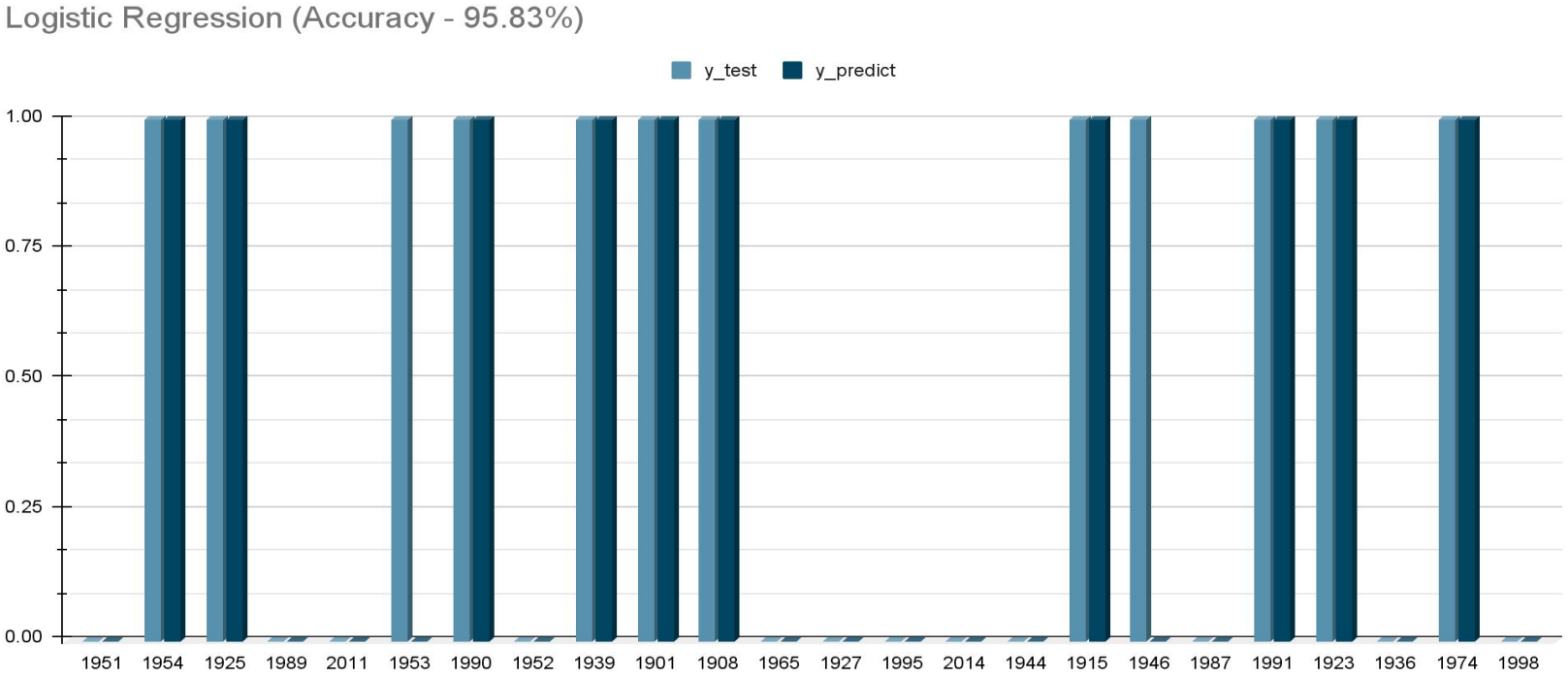
KNN Classifier (Accurasy - 75%)



Logistic Regression

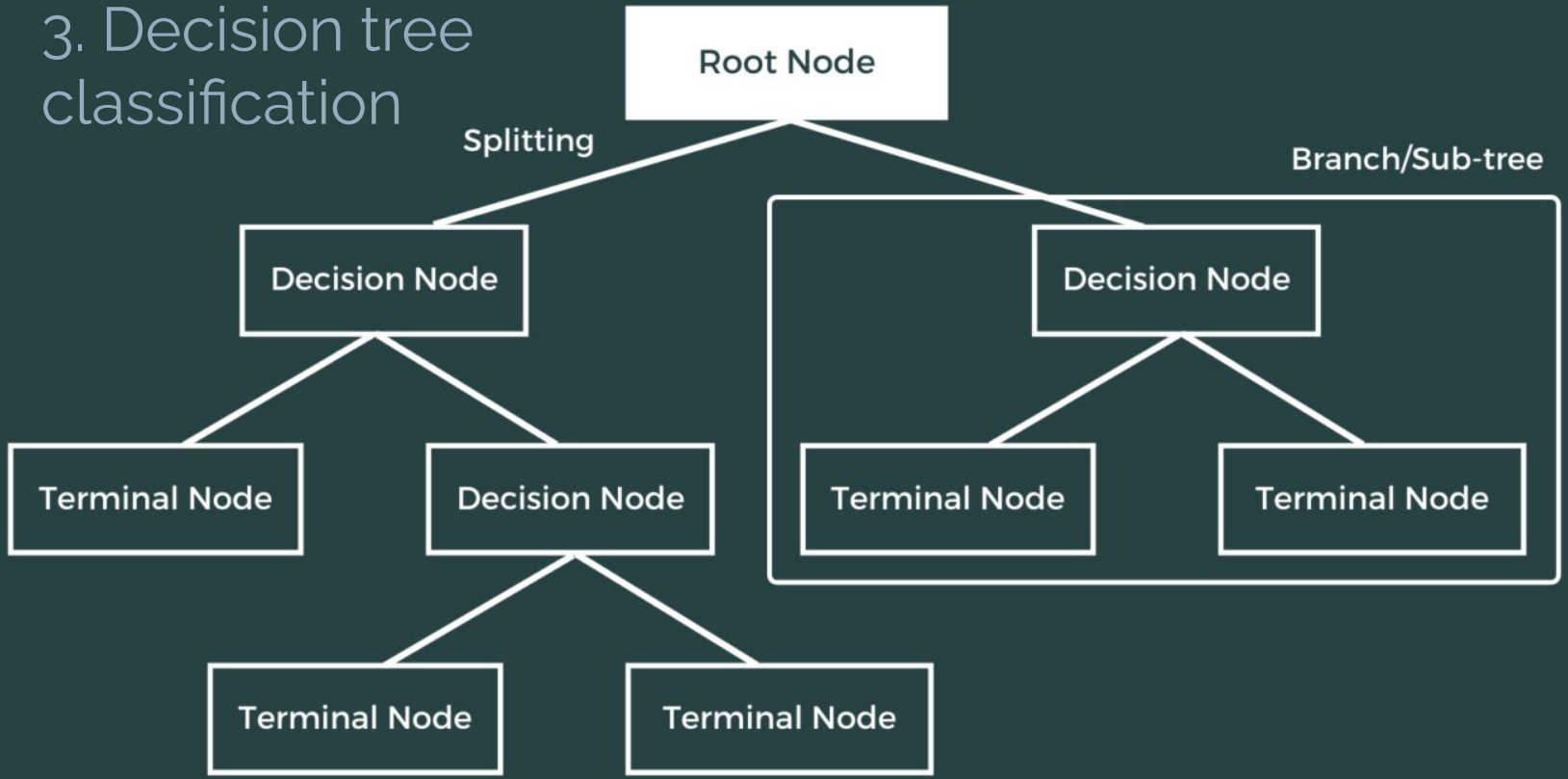


https://www.theclickreader.com/wp-content/uploads/2020/08/Logistic-Regression-1536x864.png

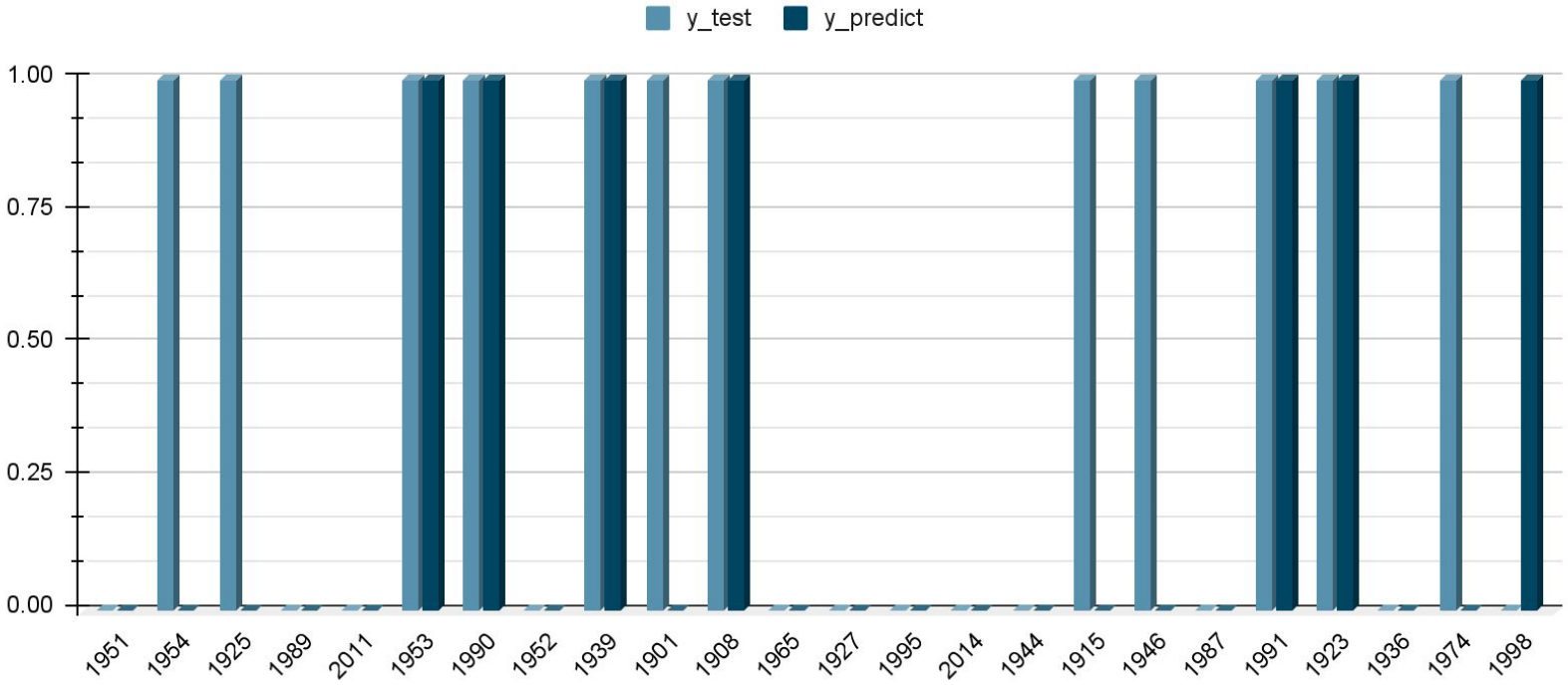

Logistic Regression (Accuracy - 95.83%)
y_test
y_predict
1.00
0.75
0.50
0.25
0.00
1951
1954
1925
1989
2011
1953
1990
1952
1939
1901
1908
1965
1927
1995
2014
1944
1915
1946
1987
1991
1923
1936
1974
1998

Decision Tree Process

3. Decision tree classification

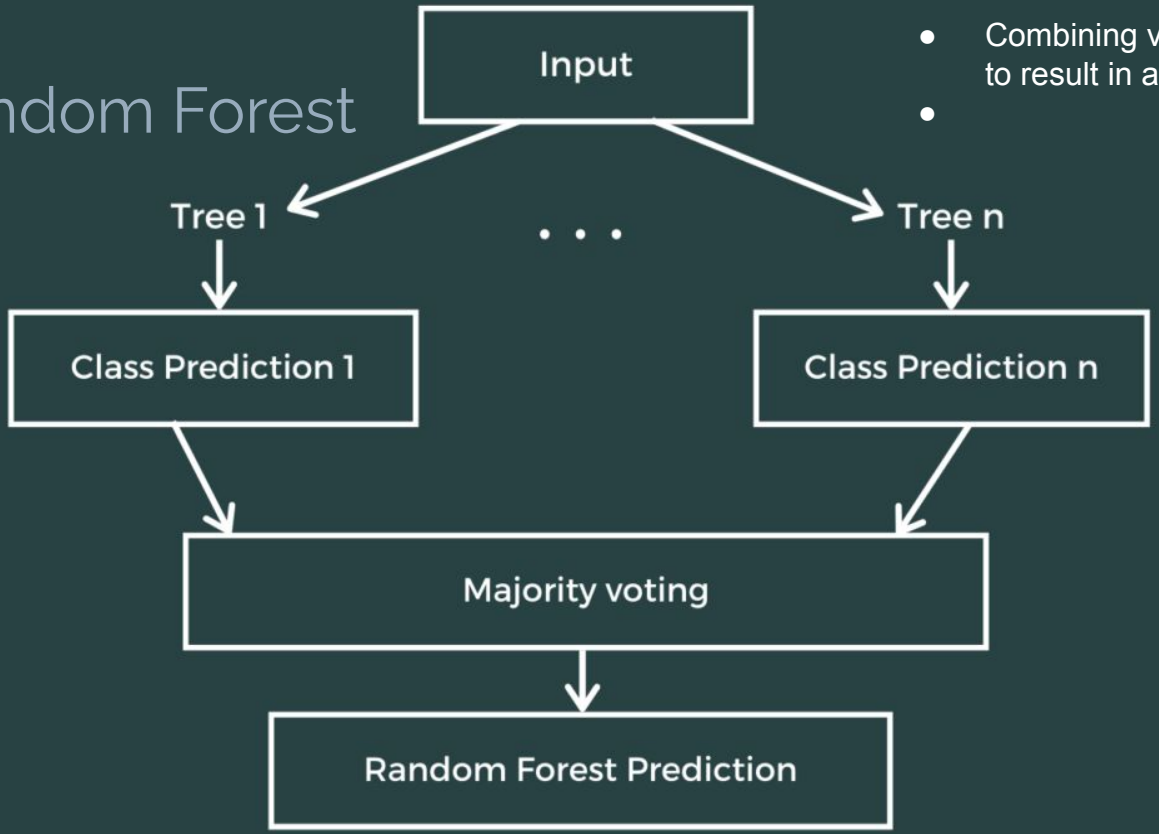


Decision tree classification (Accuracy - 79.17%)



Random Forest Process

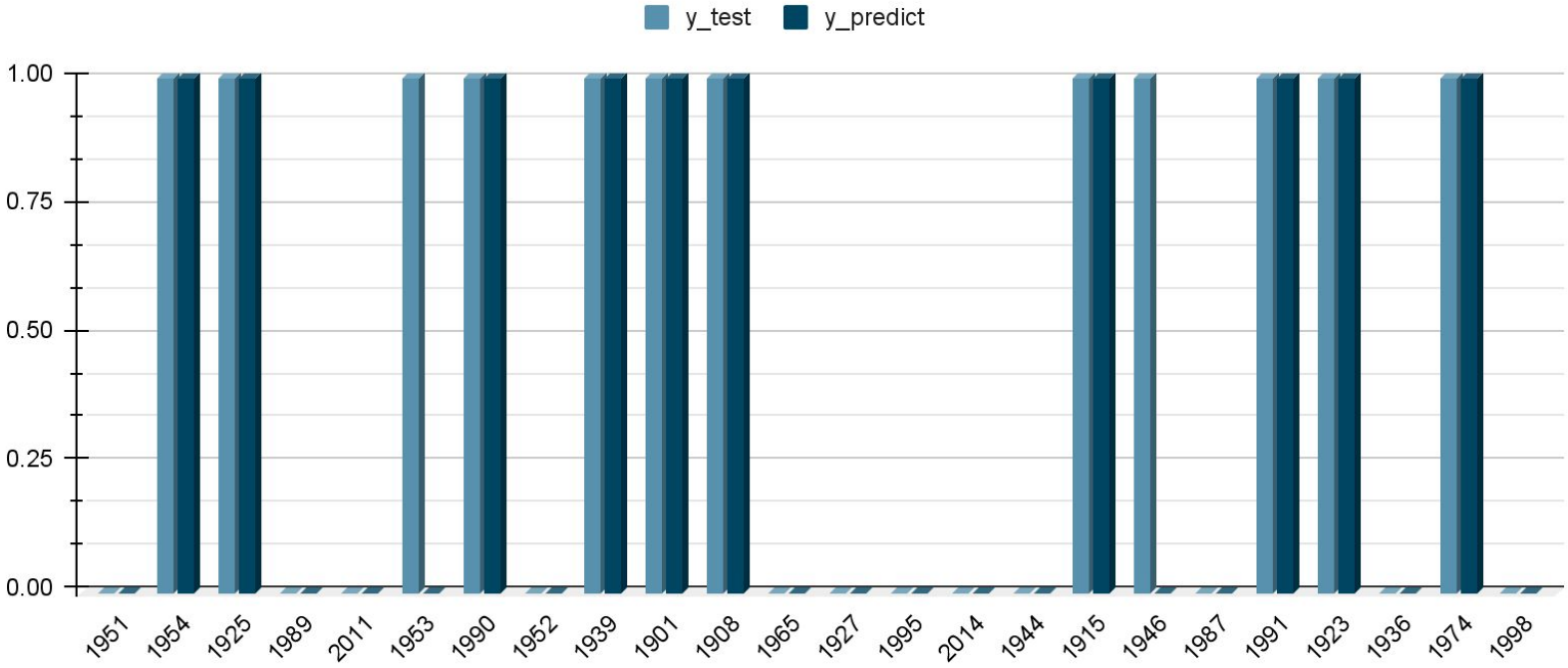
4. Random Forest



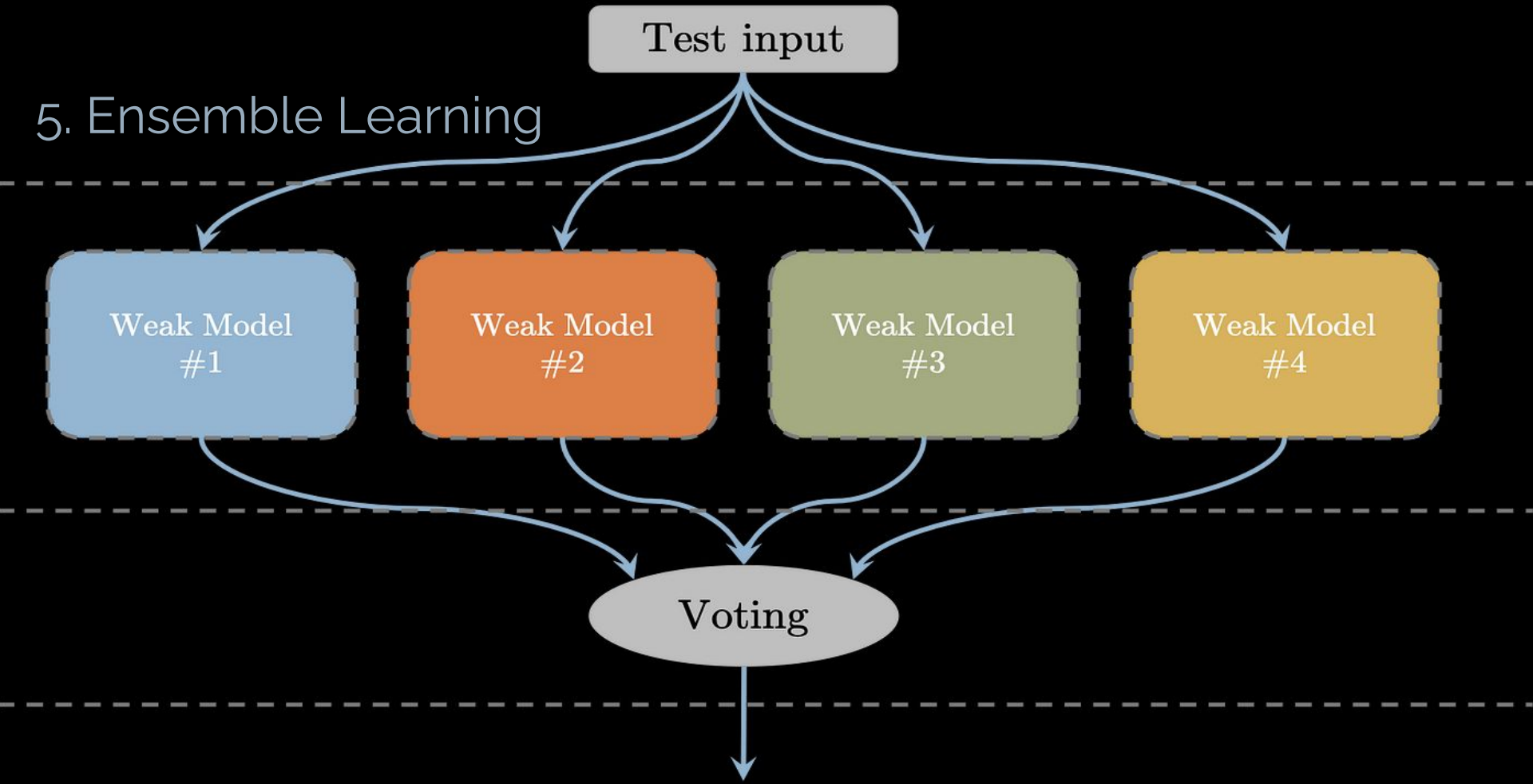
- Combining various decision trees to result in a final class prediction.
-

https://www.theclickreader.com/wp-content/uploads/2020/07/rf-classifier-1536x864.png

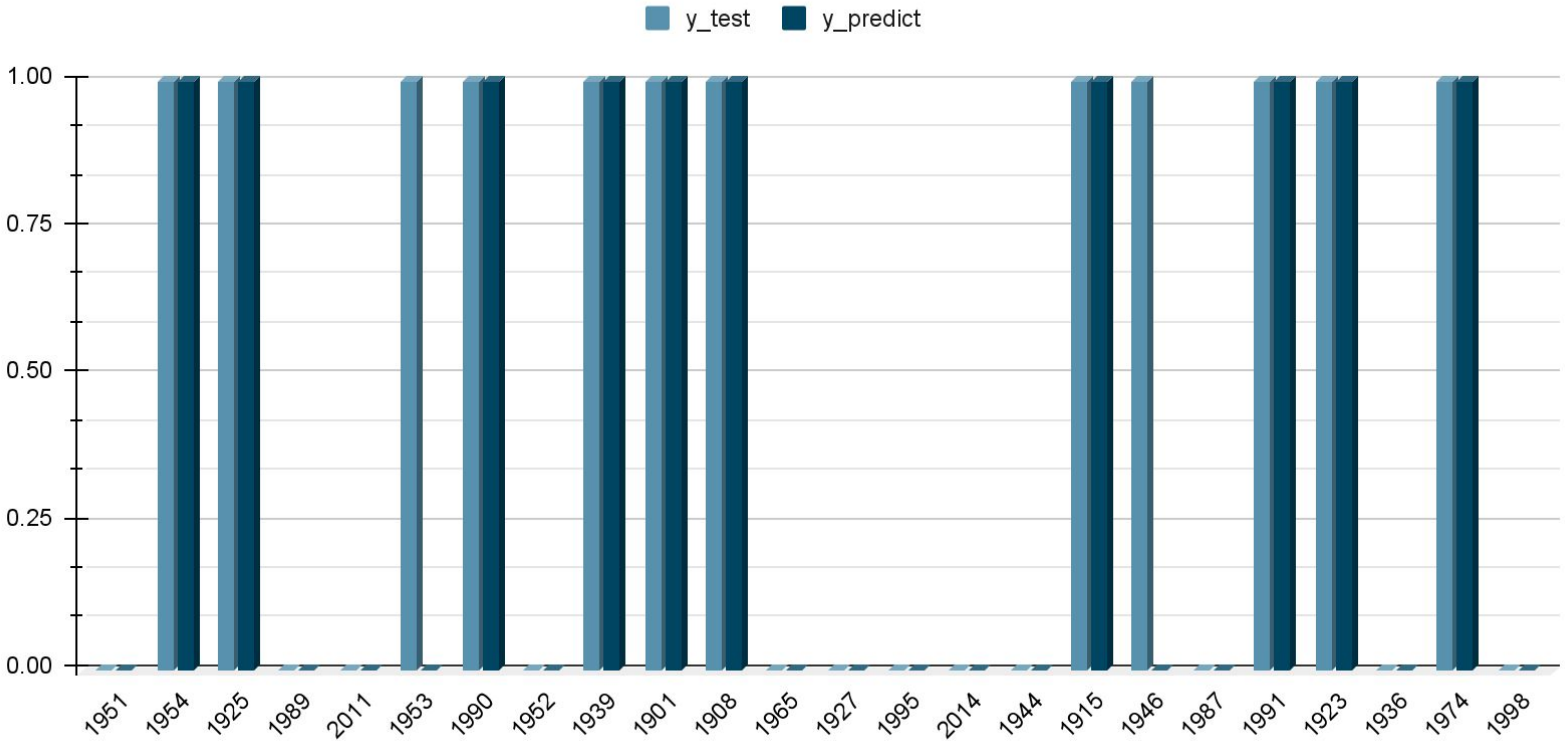
Random Forest (Accuracy - 87.5%)



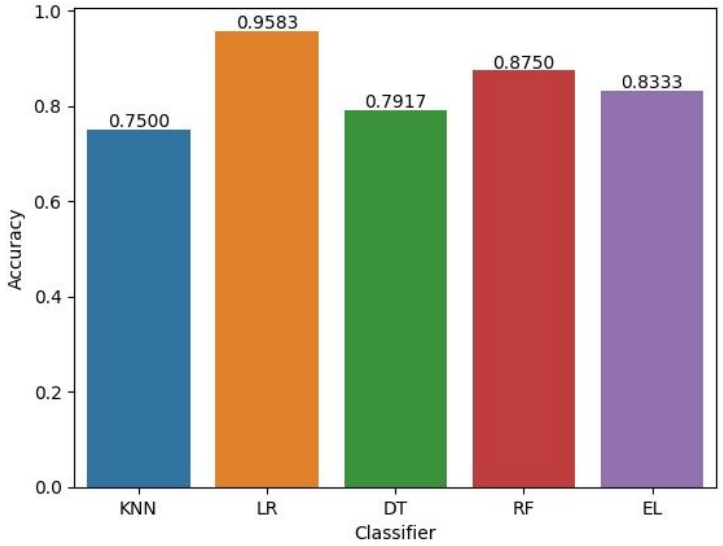
5. Ensemble Learning



Ensemble Learning (Accuracy - 83.33%)



Comparing all Models



- Among all the implemented ML algorithms we can see that *Logistic Regression* stands out with a better accuracy score of 95.83%
- So, we can choose this *Logistic Regression* model to predict the risk of floods.

Code_Notebook(here you can find the code)

Conclusion:

Logistic regression is a probabilistic model that estimates the probability of a data point belonging to a particular class. This makes it well-suited for problems where the goal is to estimate the probability of a certain event occurring, such as in medical diagnosis or credit risk assessment.

So for any such risk predictions logistic regression can be well suited which we have also seen in our observations.

Division of work:

Sana- Data scraping, Implementation of KNN,logistic regression models

Avinash- Implemented decision tree classification, random forest classification models

Vamsidhar- Implemented ensemble learning model and wrote a code to compare models with the found dataset.

Git link:https://github.com/sana-md/DM_Project

Literature Review

Floods are one of the most common and devastating natural disasters that affect people all around the world. Flood risk management involves various processes such as flood forecasting, warning, evacuation planning, and infrastructure protection.

In recent years, artificial intelligence (AI) and machine learning (ML) techniques have emerged as promising tools for improving flood risk management. In this literature review, we will explore the recent advances in flood risk management using AI and ML models.

Flood forecasting:

Flood forecasting is an essential component of flood risk management, as it helps in predicting the occurrence and severity of floods. AI and ML techniques have been used extensively for improving flood forecasting.

Flood warning:

Flood warning systems are designed to provide timely and accurate information about the occurrence and severity of floods. AI and ML techniques have been used for improving flood warning systems.

References

- https://data.gov.in/resource/rainfall-north-east-india-and-its-departure-normal-monsoon-session-1901-2019
- https://medium.com/@mashudumudau/flood-prediction-using-kerala-state-using-machine-learning-7cb4e1695ac8
- https://www.theclickreader.com/k-nearest-neighbours-knn-classifier/