STOCK PRICE PREDICTION

IBM NAAN MUDHALVAN

PROJECT SUBMISSION PHASE-3

Dataset is taken from kaggle competition and it can be downloaded from here:

https://www.kaggle.com/datasets/prasoonkottarathil/microsoft-lifetime-stocks-dataset

Machine Learning Algorithms:

Decision Trees and Random Forests: Useful for capturing complex relationships in the data.

Gradient Boosting Models (e.g., XGBoost, LightGBM): Excellent for predictive modeling and handling non-linear relationships.

Predicting Stock Price involves using data science techniques. A simple algorithmic approach could involve:

Scikit-Research:

A broadly used python library that gives various feature selection, extraction, and preprocessing tools. It provides a steady API, making enforcing numerous feature engineering strategies easy. Its wide adoption guarentees tremendous community suppoprt and resources.

Applications: Handling missing values, transforming categorical variables using one-hot encoding, and standardizing features with scaling strategies.

Scikit-Learn Models:

```
from sklearn.preprocessing import MinMaxScaler
from sklearn.linear_model import
LinearRegression
from sklearn.metrics importmean_squared_error,mean_absolute_error, r2_score
from sklearn.ensemble import
RandomForestRegress
from xgboost.sklearn import XGBRegressor
from sklearn.model_selection import KFold, cross_val_score, train_test_split
```

Data set: we provide the screenshots of our data set which is downloaded here:https://www.kaggle.com/datasets/prasoonkottarathil/microsoft-lifetime-stocks-dataset

	A	B	C	D	E	F	G
1	TV	Radio	Newspape	Sales			
2	230.1	37.8	69.2	22.1			
3	44.5	39.3	45.1	10.4			
4	17.2	45.9	69.3	12			
5	151.5	41.3	58.5	16.5			
6	180.8	10.8	58.4	17.9			
7	8.7	48.9	75	7.2			
8	57.5	32.8	23.5	11.8			
9	120.2	19.6	11.6	13.2			
10	8.6	2.1	1	4.8			
11	199.8	2.6	21.2	15.6			
12	66.1	5.8	24.2	12.6			
13	214.7	24	4	17.4			
14	23.8	35.1	65.9	9.2			
15	97.5	7.6	7.2	13.7			
16	204.1	32.9	46	19			
17	195.4	47.7	52.9	22.4			
18	67.8	36.6	114	12.5			
19	281.4	39.6	55.8	24.4			
20	69.2	20.5	18.3	11.3			
21	147.3	23.9	19.1	14.6			
22	218.4	27.7	53.4	18			
23	237.4	5.1	23.5	17.5			
24	13.2	15.9	49.6	5.6			
25	228.3	16.9	26.2	20.5			
26	62.3	12.6	18.3	9.7			
27	262.9	3.5	19.5	17			
28	142.9	29.3	12.6	15			
29	240.1	16.7	22.9	20.9			
30	248.8	27.1	22.9	18.9			
31	70.6	16	40.8	10.5			
32	292.9	28.3	43.2	21.4			
33	112.9	17.4	38.6	11.9			
34	97.2	1.5	30	13.2			
35	265.6	20	0.3	17.4			
36	95.7	1.4	7.4	11.9			
37	290.7	4.1	8.5	17.8			

hg

38	266.9	43.8	5	25.4			
39	74.7	49.4	45.7	14.7			
40	43.1	26.7	35.1	10.1			
41	228	37.7	32	21.5			
42	202.5	22.3	31.6	16.6			
43	177	33.4	38.7	17.1			
44	293.6	27.7	1.8	20.7			
45	206.9	8.4	26.4	17.9			
46	25.1	25.7	43.3	8.5			
47	175.1	22.5	31.5	16.1			
48	89.7	9.9	35.7	10.6			
49	239.9	41.5	18.5	23.2			
50	227.2	15.8	49.9	19.8			
51	66.9	11.7	36.8	9.7			
52	199.8	3.1	34.6	16.4			
53	100.4	9.6	3.6	10.7			
54	216.4	41.7	39.6	22.6			
55	182.6	46.2	58.7	21.2			
56	262.7	28.8	15.9	20.2			

56	262.7	28.8	15.9	20.2
57	198.9	49.4	60	23.7
58	7.3	28.1	41.4	5.5
59	136.2	19.2	16.6	13.2
60	210.8	49.6	37.7	23.8
61	210.7	29.5	9.3	18.4
62	53.5	2	21.4	8.1
63	261.3	42.7	54.7	24.2
64	239.3	15.5	27.3	20.7
65	102.7	29.6	8.4	14
66	131.1	42.8	28.9	16
67	69	9.3	0.9	11.3
68	31.5	24.6	2.2	11
69	139.3	14.5	10.2	13.4
70	237.4	27.5	11	18.9
71	216.8	43.9	27.2	22.3
72	199.1	30.6	38.7	18.3
73	109.8	14.3	31.7	12.4

74	26.8	33	19.3	8.8
75	129.4	5.7	31.3	11
76	213.4	24.6	13.1	17
77	16.9	43.7	89.4	8.7
78	27.5	1.6	20.7	6.9
79	120.5	28.5	14.2	14.2
80	5.4	29.9	9.4	5.3
81	116	7.7	23.1	11
82	76.4	26.7	22.3	11.8
83	239.8	4.1	36.9	17.3
84	75.3	20.3	32.5	11.3
85	68.4	44.5	35.6	13.6
86	213.5	43	33.8	21.7
87	193.2	18.4	65.7	20.2
88	76.3	27.5	16	12
89	110.7	40.6	63.2	16
90	88.3	25.5	73.4	12.9
91	109.8	47.8	51.4	16.7

92	134.3	4.9	9.3	14
93	28.6	1.5	33	7.3
94	217.7	33.5	59	19.4
95	250.9	36.5	72.3	22.2
96	107.4	14	10.9	11.5
97	163.3	31.6	52.9	16.9
98	197.6	3.5	5.9	16.7
99	184.9	21	22	20.5
100	289.7	42.3	51.2	25.4
101	135.2	41.7	45.9	17.2
102	222.4	4.3	49.8	16.7
103	296.4	36.3	100.9	23.8
104	280.2	10.1	21.4	19.8
105	187.9	17.2	17.9	19.7
106	238.2	34.3	5.3	20.7
107	137.9	46.4	59	15
108	25	11	29.7	7.2
109	90.4	0.3	23.2	12
110	13.1	0.4	25.6	5.3
111	255.4	26.9	5.5	19.8
112	225.8	8.2	56.5	18.4
113	241.7	38	23.2	21.8
114	175.7	15.4	2.4	17.1
115	209.6	20.6	10.7	20.9
116	78.2	46.8	34.5	14.6
117	75.1	35	52.7	12.6
118	139.2	14.3	25.6	12.2
119	76.4	0.8	14.8	9.4
120	125.7	36.9	79.2	15.9
121	19.4	16	22.3	6.6
122	141.3	26.8	46.2	15.5
123	18.8	21.7	50.4	7
124	224	2.4	15.6	16.6
125	123.1	34.6	12.4	15.2
126	229.5	32.3	74.2	19.7
127	87.2	11.8	25.9	10.6

128	7.8	38.9	50.6	6.6		
129	80.2	0	9.2	11.9		
130	220.3	49	3.2	24.7		
131	59.6	12	43.1	9.7		
132	0.7	39.6	8.7	1.6		
133	265.2	2.9	43	17.7		
134	8.4	27.2	2.1	5.7		
135	219.8	33.5	45.1	19.6		
136	36.9	38.6	65.6	10.8		
137	48.3	47	8.5	11.6		
138	25.6	39	9.3	9.5		
139	273.7	28.9	59.7	20.8		
140	43	25.9	20.5	9.6		
141	184.9	43.9	1.7	20.7		
142	73.4	17	12.9	10.9		
143	193.7	35.4	75.6	19.2		
144	220.5	33.2	37.9	20.1		
145	104.6	5.7	34.4	10.4		
146	96.2	14.8	38.9	12.2		

146	96.2	14.8	38.9	12.3
147	140.3	1.9	9	10.3
148	240.1	7.3	8.7	18.2
149	243.2	49	44.3	25.4
150	38	40.3	11.9	10.9
151	44.7	25.8	20.6	10.1
152	280.7	13.9	37	16.1
153	121	8.4	48.7	11.6
154	197.6	23.3	14.2	16.6
155	171.3	39.7	37.7	16
156	187.8	21.1	9.5	20.6
157	4.1	11.6	5.7	3.2
158	93.9	43.5	50.5	15.3
159	149.8	1.3	24.3	10.1
160	11.7	36.9	45.2	7.3
161	131.7	18.4	34.6	12.9
162	172.5	18.1	30.7	16.4
163	85.7	35.8	49.3	13.3
164	188.4	18.1	25.6	19.9
165	163.5	36.8	7.4	18
166	117.2	14.7	5.4	11.9
167	234.5	3.4	84.8	16.9
168	17.9	37.6	21.6	8
169	206.8	5.2	19.4	17.2
170	215.4	23.6	57.6	17.1
171	284.3	10.6	6.4	20
172	50	11.6	18.4	8.4
173	164.5	20.9	47.4	17.5
174	19.6	20.1	17	7.6
175	168.4	7.1	12.8	16.7
176	222.4	3.4	13.1	16.5
177	276.9	48.9	41.8	27
178	248.4	30.2	20.3	20.2
179	170.2	7.8	35.2	16.7
180	276.7	2.3	23.7	16.8
181	165.6	10	17.6	17.6
182	156.6	2.6	8.2	15.5

182	156.6	2.6	8.3	15.5			
183	218.5	5.4	27.4	17.2			
184	56.2	5.7	29.7	8.7			
185	287.6	43	71.8	26.2			
186	253.8	21.3	30	17.6			
187	205	45.1	19.6	22.6			
188	139.5	2.1	26.6	10.3			
189	191.1	28.7	18.2	17.3			
190	286	13.9	3.7	20.9			
191	18.7	12.1	23.4	6.7			
192	39.5	41.1	5.8	10.8			
193	75.5	10.8	6	11.9			
194	17.2	4.1	31.6	5.9			
195	166.8	42	3.6	19.6			
196	149.7	35.6	6	17.3			
197	38.2	3.7	13.8	7.6			
198	94.2	4.9	8.1	14			
199	177	9.3	6.4	14.8			
200	283.6	42	66.2	25.5			
201	232.1	8.6	8.7	18.4			

We build the **Stock Price Prediction** model by loading and preprocessing the dataset and we load the historical dataset and preprocess the data for analysis.