

Day 40

DIY

Q1. Problem Statement: Naïve Bayes Classification

Write a Python program that reads the *mobile price.csv* (provided on LMS) file into a DataFrame, by doing the given task builds good naïve Bayes models.

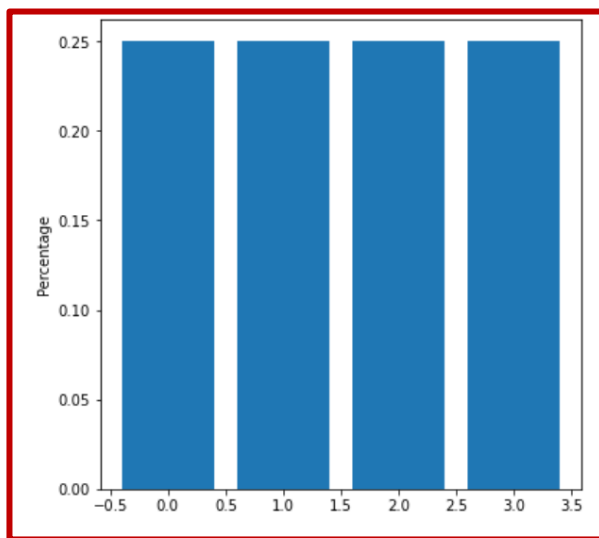
1. Load the *mobile price.csv* dataset into a DataFrame
2. Check the data type of each column and change it according to their data
3. Find missing values and treat them if you found any
4. Plot bar graph for target variable and their categories with help of seaborn and if data is not balanced then treat them using SMOT.
5. Extract independent variables (Xs) and dependent variables (Ys) into separate data objects
6. Split data into train and test DataFrame and set test size as 0.2
7. Build Gaussian and Bernoulli models, compare their accuracy

Input Table Format:

	battery_power	blue	clock_speed	dual_sim	fc	four_g	int_memory	m_dep	mobile_wt	n_cores	...	px_height	px_width	ram	sc_h	sc_w	talk_time	three_g	touch_screen	wifi	price_range
0	842	0	2.2	0	1	0	7	0.6	188	2	...	20	756	2549	9	7	19	0	0	1	1
1	1021	1	0.5	1	0	1	53	0.7	136	3	...	905	1988	2631	17	3	7	1	1	0	2
2	563	1	0.5	1	2	1	41	0.9	145	5	...	1263	1716	2603	11	2	9	1	1	0	2
3	615	1	2.5	0	0	0	10	0.8	131	6	...	1216	1786	2769	16	8	11	1	0	0	2
4	1821	1	1.2	0	13	1	44	0.6	141	2	...	1208	1212	1411	8	2	15	1	1	0	1

Sample Output:

1. Plot bar graph for target variable and their categories with help of seaborn and if data is not balanced then treat them using SMOT.



7. Build Gaussian and Bernoulli models, compare their accuracy

