Approach for the assignment

The data given was in comma seperated value format. There were two feature columns i.e. ***visitor_id*** and ***visits*** (the day of visits are in integer in a string format). As the ***visitor_id*** is unique for each visitors, so we need to convert the ***visits*** feature columns into more readable, understandable and interpretable format.

	visitor_id	visits
0	195431	5 6 10 11 31 60 79 88 105 126 127 139 141 144 ...
1	155842	9 21 37 38 40 47 51 90 110 111 118 130 132 148...
2	198112	13 27 33 41 68 73 125 130 131 144 146 152 154 ...

Firstly, I wrote few scripts that to take each row of ***visits*** column and convert in into:

1. number of times the app was opened by a particular visitor per *day of the week.*
2. number of times the app was opened by a particular visitor per *month of the year***.**
3. number of times the app was opened by a particular visitor per *year.*
4. number of times the app was opened by a particular visitor per *week of the year.*
5. average number of days between *consecutive app opening*.

Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
16.0	10.0	14.0	17.0	9.0	29.0	16.0
2.0	5.0	7.0	10.0	17.0	33.0	10.0
3.0	4.0	5.0	13.0	26.0	45.0	22.0

January	February	March	April	May	June	July	August	September	October	November	December
11.0	9.0	9.0	7.0	13.0	11.0	8.0	11.0	11.0	5.0	9.0	7.0
4.0	10.0	5.0	10.0	14.0	5.0	7.0	9.0	10.0	4.0	3.0	3.0
11.0	6.0	9.0	9.0	11.0	15.0	12.0	13.0	11.0	6.0	7.0	8.0

2018	2019	2020
40	36	35
30	24	30
36	42	40

week-00	week-01	week-02	week-03	week-04	week-05	week-06	week-07	week-08	week-09	...
1.0	4.0	3.0	1.0	2.0	3.0	1.0	5.0	NaN	2.0	...
1.0	NaN	1.0	2.0	1.0	1.0	4.0	3.0	2.0	2.0	...
1.0	2.0	5.0	2.0	1.0	2.0	2.0	NaN	4.0	1.0	...

Average_app_open_days
9
12
8

How the Prediction works:

For predicting the next day of visit for each visitor, I wrote a little script that will return the name of the *weekday* and name of the *week number* that they visited the app most, for example, 1st visitor visited the app most on ***Saturday*** and on ***week-07.***

	visitor_id	Weekdays	Week_number
0	195431	Saturday	week-07
1	155842	Saturday	week-19
2	198112	Saturday	week-25
3	72900	Saturday	week-01
4	159661	Friday	week-30

Drawbacks

The only drawbacks of this approach will be the time complexity of the scripts which takes the all the values of a particular visitor and convert the data into meaningful format.