PREDICTION

PROJECT

ESILV - School Project

HELALI NADER & EMMUNGIL YANIS

DATA SET

ONLINE SHOPPER INTENTION

	Administrative	Administrative_Duration	Informational	Informational_Duration	ProductRelated	ProductRelated_Duration	BounceRates	ExitRates
4584	0	0.000000	0	0.0	68	3177.791667	0.005970	0.011488
7409	0	0.000000	0	0.0	2	0.000000	0.200000	0.200000
9972	0	0.000000	0	0.0	63	1361.203175	0.000000	0.014762
8146	7	60.291667	0	0.0	8	108.500000	0.013333	0.048667

PageValues	SpecialDay	Month	Operating Systems	Browser	Region	Traffic Type	VisitorType	Weekend	Revenue
0.0	0.0	May	2	4	4	4	Returning_Visitor	False	False
0.0	0.0	Jul	3	2	1	1	Returning_Visitor	False	False
0.0	0.0	Nov	2	2	7	2	Returning_Visitor	False	False
0.0	0.0	Dec	2	2	2	6	Returning_Visitor	False	False

FEATURES

- Administrative: This is the number of pages of this type (administrative) that the user visisted.
- Administrative_Duration: This is the amount of time spent in this category of pages

- Informational: This is the number of pages of this type (informational) that the user visited.
- Informational_Duration: This is the amount of time spent in this category of pages.

- Product Related: This is the number of pages of this type (product related) that the user visited.
- Product Related Duration: This is the amount of time spent in this category of pages

- Bounce Rate is the percentage of visitors who enter the site from that page and then leave ("bounce") without triggering any other requests to the analytics server during that session.
- Exit Rate is calculated as for all pageviews to the page, the percentage that were the last in the session.
- Page Value represents the average value for a web page that a user visited before completing an e-commerce transaction.
- Special Day indicates the closeness of the site visiting time to a specific special day (e.g. Mother's Day, Valentine's Day) in which the sessions are more likely to be finalized with transaction.

- The dataset also includes operating system, browser, region, traffic type, visitor type as returning or new visitor, a Boolean value indicating whether the date of the visit is weekend, and month of the year.

OBJECTIVE

After the explanation of our dataset, we can define our Problematic.

Make a prediction of the user comportment is our objective here. And predict if the user will generate some revenue for the online shop or not.

TABLE OF CONTENTS
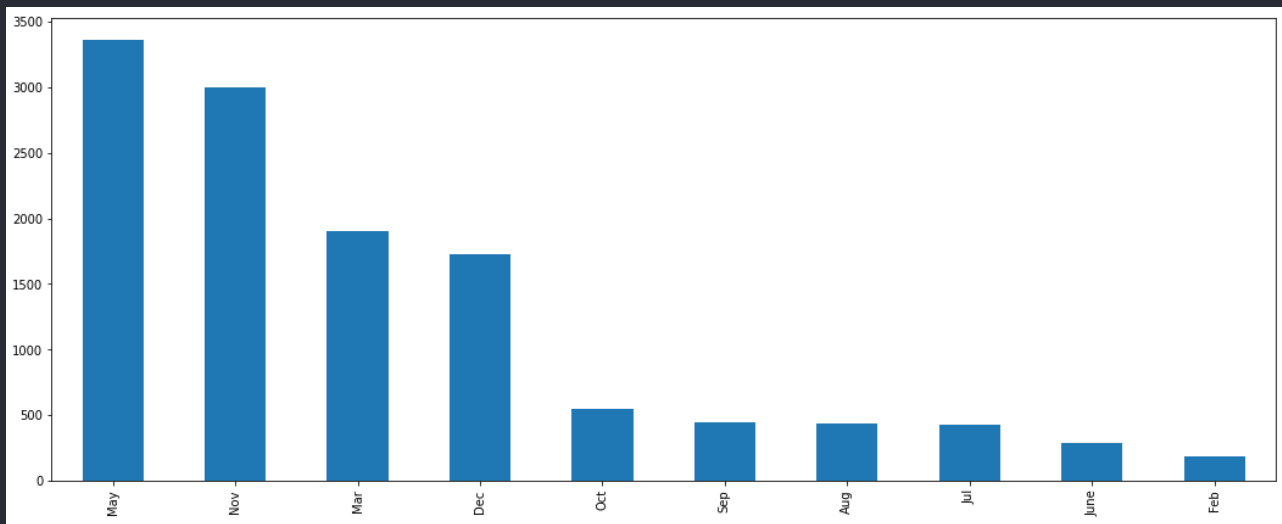
1. DATA VIZUALIZATION

2. PREPROCESSING
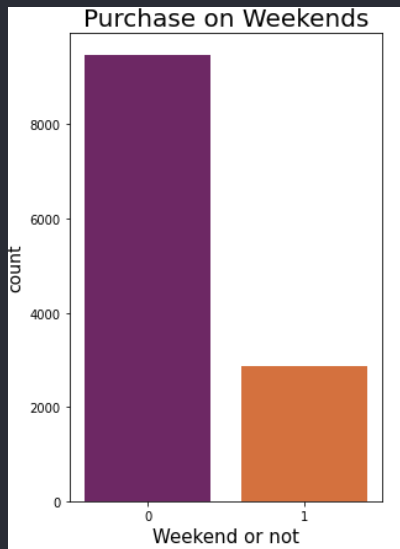
3. MODELING

4. FLASK API

DATA VIZUALIZATION

THE NUMBER OF VISITOR BY MONTH :
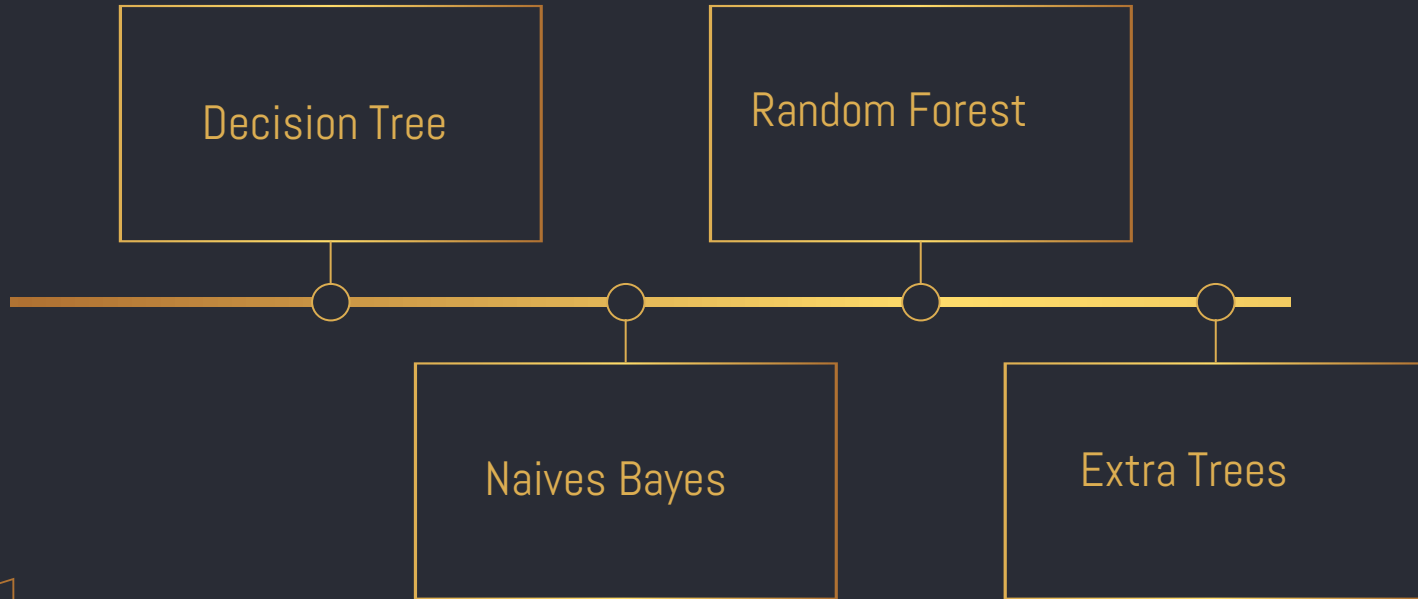

FIND ALL OF THEM ON OUR NOTEBOOK

DATA VIZUALIZATION

THE NUMBER OF PURCHASE DURING THE WEEK AND THE WEEK-END:

Accuracy

	Decision Tree	Naive Bayes	Random Forest	Extra Trees
Accuracy (%)	85,32	81,54	91	87,92

API

- Implentention of our best model in a Flask Api :

  - That permits to the user to enter the variable and he returns the prediction

THANKS!

ESILV DIA 3

HELALI Nader
EMMUNGIL Yanis