CHEE JUN HONG - PROJECT MANAGER

Bank Marketing

VLC-SGITP25-21-0659



THESE ARE THE BROAD TOPICS THIS MEETING WILL COVER.

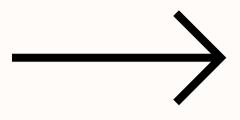
Discussion

Points of 02

> Results 03

01

04



Introduction

Data

Conclusion

Introduction

Two major objectives we're focusing

STREAMLING AND OPTIMIZING 01 **EFFICIENCY AND RESOURCES**

With the introduction of new Machine Learning team, we can better optimize our budget and manpower by identifying customers with higher potential to sign on our services offered.

02



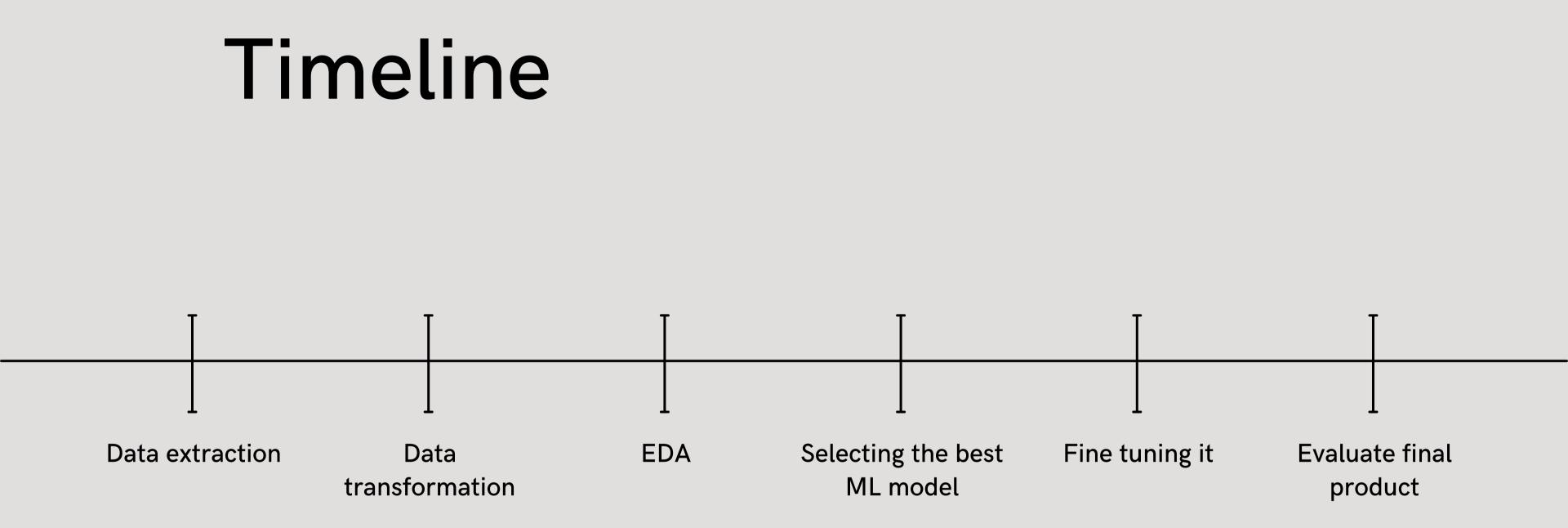
INCREASE OUR SALES REVENUE

Contact and engage with these customers and also upsell our current services.



\$65,000,000

For a typical Fortune 1000 company, just a 10% increase in data accessibility will result in more than \$65 million additional net income. - Bernard Marr. Forbes

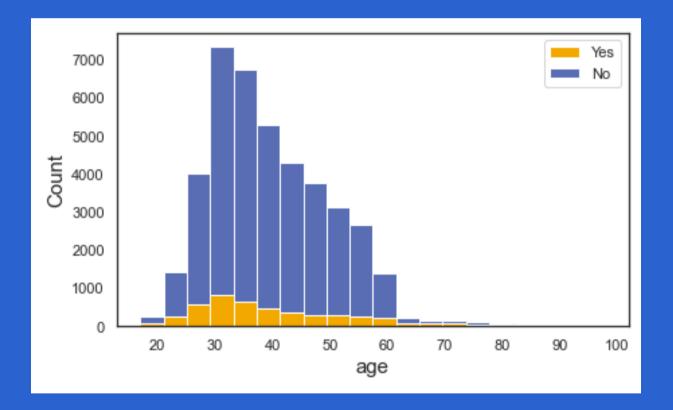




700%

The statistics show the banking sector generates unparalleled quantities of data. The amount of data generated each second in the financial industry will grow 700% in 2021. - Madeline Connall - Sigma

Data Information



The info of our Data selected

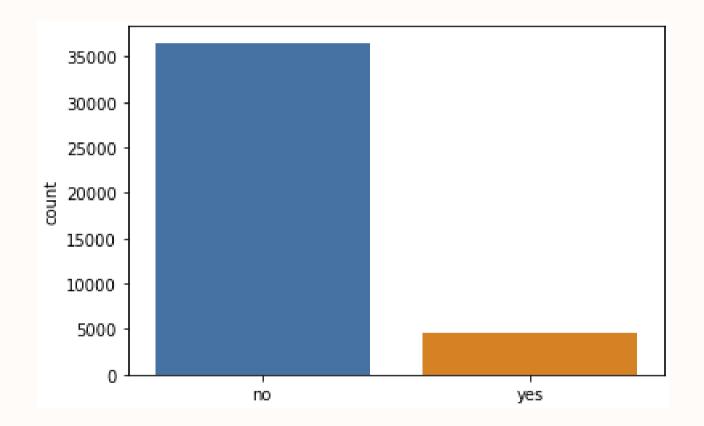
df.info()

<class 'pandas.core.frame.dataframe'=""></class>					
RangeIndex: 41188 entries, 0 to 41187 Data columns (total 21 columns):					
			•	D ±	
#	Column	Non-NI	ull Count	Dtype	
0	age	41188	non-null	int64	
1	job	41188	non-null	object	
2	marital		non-null	object	
3	education	41188	non-null	object	
4	default	41188	non-null	object	
5	housing	41188	non-null	object	
6	loan	41188	non-null	object	
7	contact	41188	non-null	object	
8	month	41188	non-null	object	
9	day_of_week	41188	non-null	object	
10	duration	41188	non-null	int64	
11	campaign	41188	non-null	int64	
12	pdays	41188	non-null	int64	
13	previous	41188	non-null	int64	
14	poutcome	41188	non-null	object	
15	emp.var.rate	41188	non-null	float64	
16	cons.price.idx	41188	non-null	float64	
17	<pre>cons.conf.idx</pre>	41188	non-null	float64	
18	euribor3m	41188	non-null	float64	
19	nr.employed	41188	non-null	float64	
20	У	41188	non-null	object	
dtypes: float64(5), int64(5), object(11)					
memory usage: 6.6+ MB					

df.shape

(41188, 21)

Desired Target



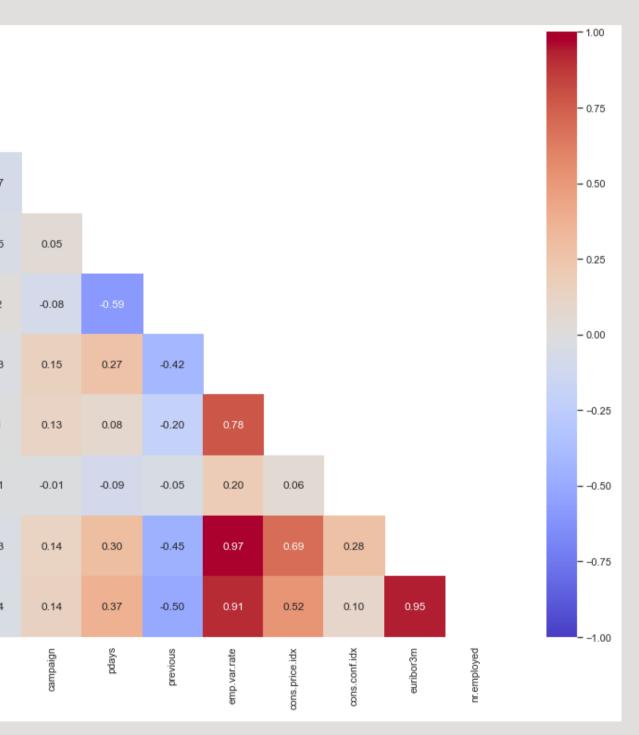
Has the client subscribed a term deposit?

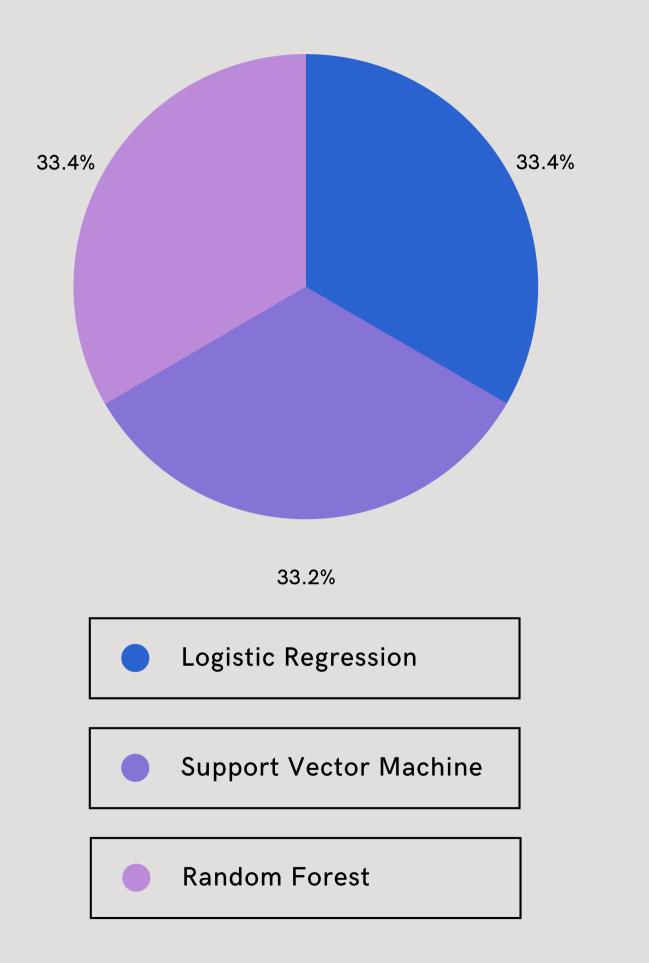
	- duration - euribor3m							-
	age nr.employed campaign pdays							
	campaign -							
	poutcome success							
	poutcome_success - cons.conf.idx -							
	emp.var.rate cons.price.idx housing_yes							
	previous -							
	marital married							
	day of week wed - day of week mon -							
	day_of_week_thu - day_of_week_tue -							
	oducation Technician							
	education_university.degree marital_married day of week wed day of week wed day of week thu day of week thu day of week thu bb technician education_high.school marital_single							
feature_names	contact telephone education_professional.course job_blue-collar default_unknown							
an	job blue-collar							
5	education basic 9y - poutcome_nonexistent -							
ure	job management							
eat	job services							
ļ	job_management job_services job_services job_retired month_oct education_unknown month_mar							
	month_mar							
	montn_mar month_may education_bastc.6y job_self-employed job_entrepreneur job_unemployed job_housemaid job_housemaid month_jul month_jul							
	job_self-employed -							
	job_entrepreneur ·	-						
	job_housemaid -							
	month jul -	1						
	loan unknown - housing_unknown - month_aug - month_nov -							
	month_aug	Ē						
	IOD UNKNOWN -							
	month_sep	1						
	marital unknown -	•						
	education illiterate default_yes							
	0.	00 0.05	0.10	0.1	5	0.20	0.25	0.30
				feature_i	mportance			
				=				

Correlation Coefficient

High relationship between 3 months Euribor, the quarterly indicator of employment and number of employees

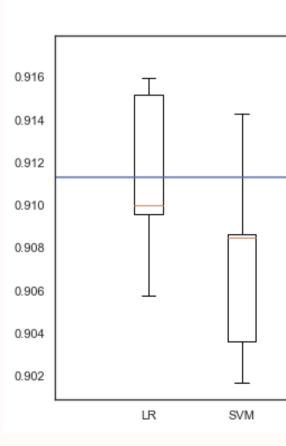
age		
duration	-0.00	
campaign	0.00	-0.07
pdays	-0.03	-0.05
previous	0.02	0.02
emp.var.rate	-0.00	-0.03
cons.price.idx	0.00	0.01
cons.conf.idx	0.13	-0.01
euribor3m	0.01	-0.03
nr.employed	-0.02	-0.04
	age	duration





Models Selection

LR : Mean score 0.9113 (Std 0.0038) SVM: Mean score 0.9073 (Std 0.0044) RF : Mean score 0.9128 (Std 0.0039) Wall time: 2min 10s



Algorithm Comparison

Model Improvement

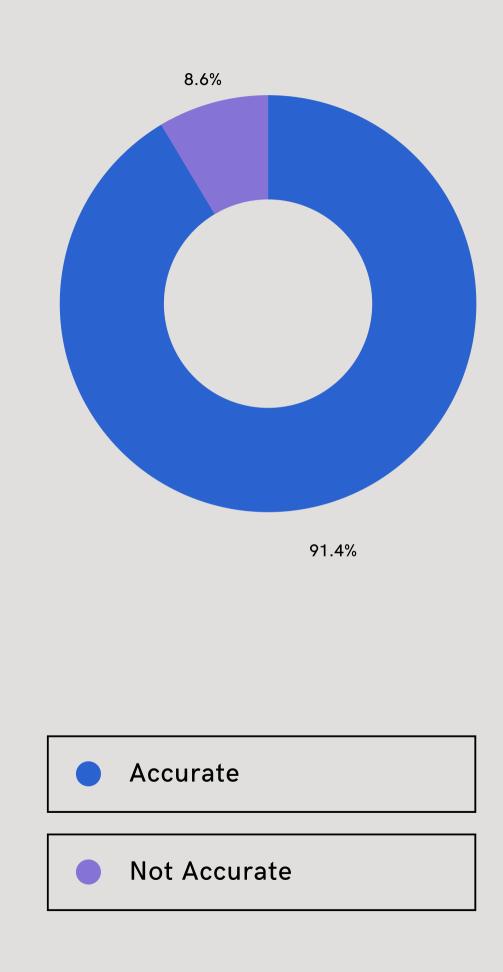
Random Forest



RandomForestClassifier(min_samples_split=10, n_estimators=1000, n_jobs=-1, oob_score=True, random_state=41) {'max_depth': None, 'min_samples_split': 10, 'n_estimators': 1000} 0.9156904400606981 Wall time: 15min 54s

Results

Final Model (RandomForestClassifier) Metrics -Accuracy : 0.9141781985918912 Precision: 0.6687022900763359 Recall : 0.47198275862068967 F1 Score : 0.5533796588755528 precision recall f1-score support 0 0.94 0.97 0.95 7310 1 0.67 0.47 0.55 928 0.91 8238 accuracy macro avg 0.75 0.80 0.72 8238 weighted avg 0.91 0.91 0.91 8238



Conclusion

01 GOOD ACCURACY

02 COMFORTABLE FALSE POSITIVE

03 USEFULLINESS MOVING FORWARD With over 91% accuracy, we will be more informed on our targeted customers.

We would rather not pass up on the opportunity to get a customer onboard even with the slightest potential.

Utilizing this Machine Learning on dataset will improve everything from optimized resource allocation which will in turn increase our revenue.

Recommendations

INCREASE BRAND AND PRODUCT AWARENESS

Allocate more budget through traditional marketing and increase our digital presence. This is to engage the general population and let them feel aware and at the same time, more comfortable with our Bank....

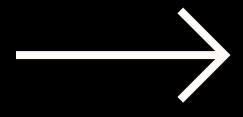
STRECH THE LIMIT

Focus on customer retention, upselling and refinancing.



Future Opportunities

EXPANSION	COLLABOR
To utilize data on more	Work with influ
department. Eg, processing of job	presence with
applications and cover letters.	class.



RATION

luences. More the young working

THANK YOU!



Bernard Marr. (2015). Forbes https://www.forbes.com/sites/bernardmarr/2015/09/30/big-data-20mind-boggling-facts-everyone-must-read/?sh=2a32576b17b1

Resource Page

Dataset from UCI Machine Learning Repository

https://archive.ics.uci.edu/ml/datasets/bank+marketing

Presentation template from Canva

Blue and Gray Swiss Gym Business Meeting Visual Charts Presentation

Madeline Connall. (2020). Sigma

https://www.sigmacomputing.com/blog/top-20-big-data-statistics/