

# Assignment 1 Extension: Streamlit Deployment

## Data Sample Introduction

We've completed the extension of the Assignment 1 and have our Streamlit app ready! Sample data points have been created to test our app with, we need to make sure that the outputs we get from the app and from our notebook are the same.

We have three different samples:

1. **Data Point 1:** Young professional with stable finances (likely to subscribe)
2. **Data Point 2:** Retired person with moderate finances (moderate likelihood)
3. **Data Point 3:** Student with minimal contact history (less likely)

```
# Data Point 1: Young professional with good finances – likely to subscribe
"age": 32,
"job": "management",
"marital": "single",
"education": "tertiary",
"default": "no",
"balance": 5000,
"housing": "yes",
"loan": "no",
"contact": "cellular",
"day": 15,
"month": "may",
"duration": 300,
"campaign": 2,
"pdays": 30,           # Previously contacted
"previous": 1,
"poutcome": "success",
"pdays_contacted": True, # Engineered feature
"pdays_clean": 30.0    # Engineered feature
```

```
# Data Point 2: Retired person with moderate finances – moderate likelihood
"age": 65,
"job": "retired",
"marital": "married",
"education": "secondary",
"default": "no",
"balance": 2500,
"housing": "no",
"loan": "no",
"contact": "telephone",
"day": 10,
"month": "jun",
"duration": 180,
"campaign": 3,
"pdays": -1,           # Never contacted before
"previous": 0,
"poutcome": "unknown",
"pdays_contacted": False, # Engineered feature
"pdays_clean": np.nan    # Engineered feature
```

```
# Data Point 3: Student with minimal contact – less likely
"age": 22,
"job": "student",
"marital": "single",
"education": "tertiary",
"default": "no",
"balance": 500,
"housing": "no",
"loan": "yes",
"contact": "cellular",
"day": 20,
"month": "aug",
"duration": 100,
"campaign": 5,
"pdays": -1,           # Never contacted before
"previous": 0,
"poutcome": "unknown",
"pdays_contacted": False, # Engineered feature
"pdays_clean": np.nan    # Engineered feature
```

## Notebook Test

Before working with our app, we'll perform a likelihood to subscribe test on our notebook with our top performing model.

```
=====
PREDICTION RESULTS (to compare with Streamlit)
=====

📊 Data Point 1:
Prediction: yes
Probabilities: no=22.7%, yes=77.3%

📊 Data Point 2:
Prediction: yes
Probabilities: no=25.2%, yes=74.8%


📊 Data Point 3:
Prediction: no
Probabilities: no=54.0%, yes=46.0%

=====
SUMMARY TABLE
=====
Data Point Prediction Prob(no) Prob(yes)
1          yes    22.7%    77.3%
2          yes    25.2%    74.8%
3          no     54.0%    46.0%
```

Here we have our reference predictions.


## App Introduction

Now we'll go over our Streamlit app, which has been created so that we can introduce all the sample data points manually to check how likely someone is to say yes to the bank's offer.




# Bank Term Deposit Subscription Prediction

Predict whether a customer will subscribe to a term deposit based on their profile.

 Model Test Accuracy: 0.8682 (86.82%)

## Input Customer Information

Fill in the values for each feature and click Predict to get the model's prediction.



### Numerical Features

age

32

?

balance

5000

?

day

15

?

duration

300

?

campaign

2

?

pdays

30

?

previous


1

?

pdays\_clean

30

?



### Categorical Features

job

management

?

marital

single

?

education

tertiary

?

default

no

?

housing

yes

?

loan

no

?

contact

cellular

?

month

may

?

poutcome


success

?

pdays\_contacted (was previously contacted?)

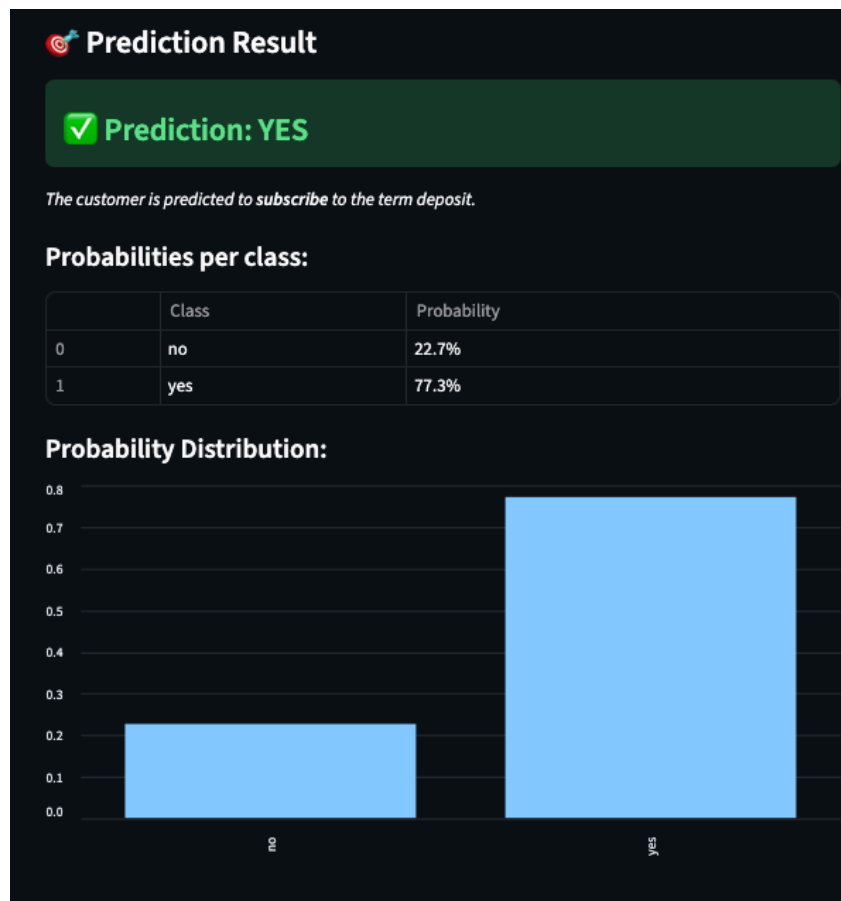
True

?

 Predict

*The app's very beautiful I might add.*

## Sample Test #1




For this example, we predict that the customer will say yes, given the 22.7/77.3 result. So we can see that it perfectly matches the probabilities provided by our model within the notebook!

```
=====
SUMMARY TABLE
=====
```

Data Point	Prediction	Prob(no)	Prob(yes)
1	yes	22.7%	77.3%
2	yes	25.2%	74.8%
3	no	54.0%	46.0%

## Sample Test #2

First we enter de sample data

 **Numerical Features**

age?  
65

balance?  
2500

day?  
10


duration?  
180

campaign?  
3

pdays?  
-1

previous?  
0

pdays\_clean?  
Leave empty if not contacted

 **Categorical Features**

job?  
retired

marital?  
married

education?  
secondary

default?  
no

housing?  
no


loan?  
no

contact?  
telephone

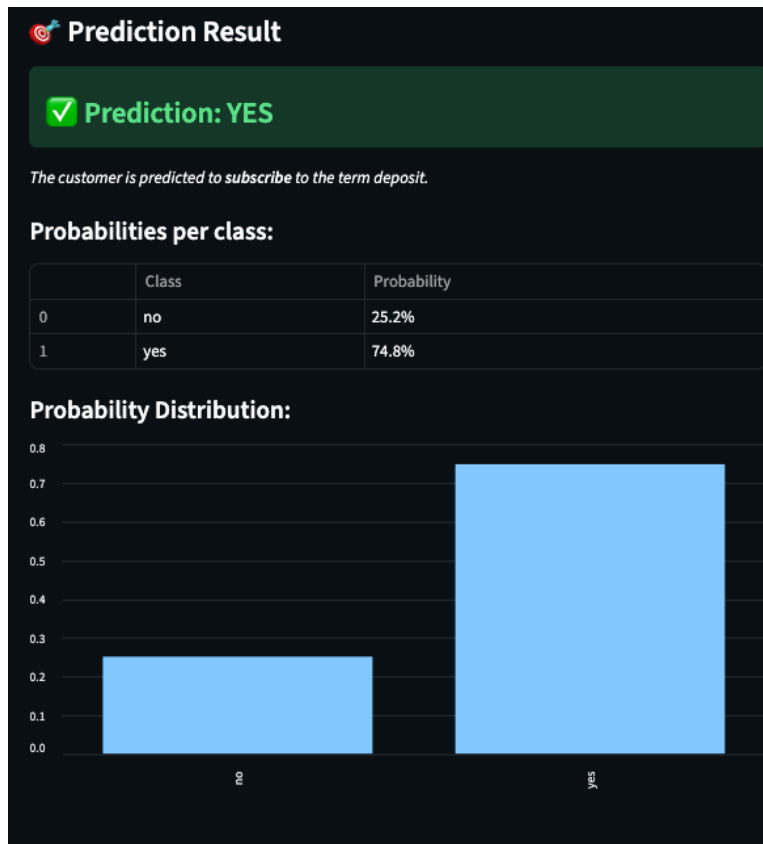
month?  
jun

poutcome?  
unknown

pdays\_contacted (was previously contacted?)?  
False

 Predict

And then we get the results:




The predictions are 25.2/74.8, so we think the customer will accept the offer. Again, we got the same results as we did with our notebook!

```
=====
SUMMARY TABLE
=====
```


Data Point	Prediction	Prob(no)	Prob(yes)
1	yes	22.7%	77.3%
2	yes	25.2%	74.8%
3	no	54.0%	46.0%

## Sample Test #3


Lastly, we'll go through our sample test #3 and introduce the sample data.

 **Numerical Features**

age	<input type="text" value="22"/>	balance	<input type="text" value="500"/>
day	<input type="text" value="20"/>	duration	<input type="text" value="100"/>
campaign	<input type="text" value="5"/>	pdays	<input type="text" value="-1"/>
previous	<input type="text" value="0"/>	pdays_clean	<input type="text" value="Leave empty if not contacted"/>

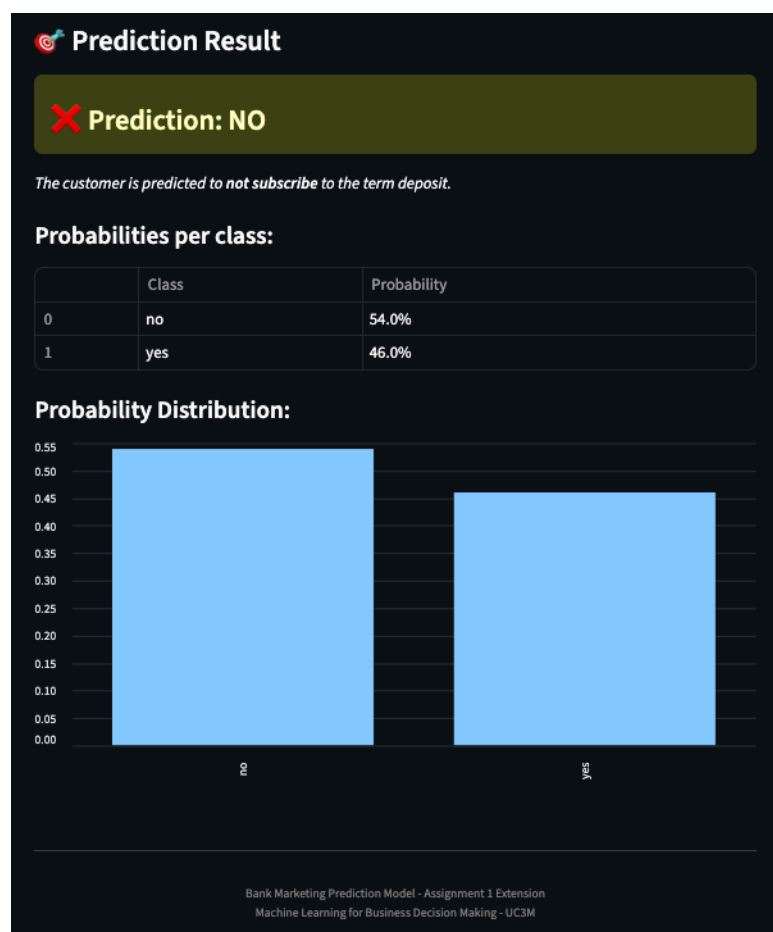
 **Categorical Features**

job	<input type="text" value="student"/>	marital	<input type="text" value="single"/>
education	<input type="text" value="tertiary"/>	default	<input type="text" value="no"/>
housing	<input type="text" value="no"/>	loan	<input type="text" value="yes"/>
contact	<input type="text" value="cellular"/>	month	<input type="text" value="aug"/>
poutcome	<input type="text" value="unknown"/>	pdays_contacted (was previously contacted?)	<input type="text" value="False"/>

 Predict

Then we click on predict and see that...





The prediction is “NO” with a 54/46 exactly matching what our notebook predicted!

SUMMARY TABLE				
Data Point	Prediction	Prob(no)	Prob(yes)	
1	yes	22.7%	77.3%	
2	yes	25.2%	74.8%	
3	no	54.0%	46.0%	