

[illegible]

	mean radius	mean texture	mean perimeter	mean area	mean smoothness	mean compactness	mean concavity	mean concave points	mean symmetry	mean fractal dimension	...	worst radius	worst texture	wor
0	17.99	10.38	122.80	1001.0	0.11840	0.27760	0.3001	0.14710	0.2419	0.07871	...	25.38	17.33	184
1	20.57	17.77	132.90	1326.0	0.08474	0.07864	0.0869	0.07017	0.1812	0.05667	...	24.99	23.41	158
2	19.69	21.25	130.00	1203.0	0.10960	0.15990	0.1974	0.12790	0.2069	0.05999	...	23.57	25.53	152
3	11.42	20.38	77.58	386.1	0.14250	0.28390	0.2414	0.10520	0.2597	0.09744	...	14.91	26.50	98
4	20.29	14.34	135.10	1297.0	0.10030	0.13280	0.1980	0.10430	0.1809	0.05883	...	22.54	16.67	152
5 rows × 30 columns														

```
data_frame['label'] = breast_cancer_dataset.target
```

```
data_frame.tail()
```

	mean radius	mean texture	mean perimeter	mean area	mean smoothness	mean compactness	mean concavity	mean concave points	mean symmetry	mean fractal dimension	...	worst texture	worst perimeter	w
564	21.56	22.39	142.00	1479.0	0.11100	0.11590	0.24390	0.13890	0.1726	0.05623	...	26.40	166.10	21
565	20.13	28.25	131.20	1261.0	0.09780	0.10340	0.14400	0.09791	0.1752	0.05533	...	38.25	155.00	17
566	16.60	28.08	108.30	858.1	0.08455	0.10230	0.09251	0.05302	0.1590	0.05648	...	34.12	126.70	17
567	20.60	29.33	140.10	1265.0	0.11780	0.27700	0.35140	0.15200	0.2397	0.07016	...	39.42	184.60	18
568	7.76	24.54	47.92	181.0	0.05263	0.04362	0.00000	0.00000	0.1587	0.05884	...	30.37	59.16	18

5 rows × 31 columns

```
data_frame.shape
```

```
(569, 31)
```

```
data_frame.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 569 entries, 0 to 568
Data columns (total 31 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   mean radius                          569 non-null    float64
1   mean texture                         569 non-null    float64
2   mean perimeter                       569 non-null    float64
3   mean area                           569 non-null    float64
4   mean smoothness                     569 non-null    float64
5   mean compactness                    569 non-null    float64
6   mean concavity                      569 non-null    float64
7   mean concave points                 569 non-null    float64
8   mean symmetry                       569 non-null    float64
9   mean fractal dimension              569 non-null    float64
10  radius error                        569 non-null    float64
11  texture error                       569 non-null    float64
12  perimeter error                     569 non-null    float64
13  area error                         569 non-null    float64
14  smoothness error                   569 non-null    float64
15  compactness error                  569 non-null    float64
16  concavity error                    569 non-null    float64
17  concave points error               569 non-null    float64
18  symmetry error                     569 non-null    float64
19  fractal dimension error            569 non-null    float64
20  worst radius                       569 non-null    float64
21  worst texture                      569 non-null    float64
22  worst perimeter                    569 non-null    float64
23  worst area                        569 non-null    float64
24  worst smoothness                   569 non-null    float64
25  worst compactness                  569 non-null    float64
26  worst concavity                    569 non-null    float64
27  worst concave points               569 non-null    float64
28  worst symmetry                     569 non-null    float64
29  worst fractal dimension            569 non-null    float64
30  label                             569 non-null    int64
dtypes: float64(30), int64(1)
memory usage: 137.9 KB
```

```
data_frame.isnull().sum()
```

```
mean radius      0
mean texture     0
mean perimeter   0
mean area        0
mean smoothness  0
mean compactness 0
mean concavity   0
mean concave points 0
mean symmetry    0
mean fractal dimension 0
radius error     0
texture error    0
perimeter error  0
area error       0
smoothness error 0
compactness error 0
concavity error  0
concave points error 0
```

```
symmetry error      0
fractal dimension error  0
worst radius        0
worst texture        0
worst perimeter      0
worst area           0
worst smoothness     0
worst compactness    0
worst concavity      0
worst concave points 0
worst symmetry       0
worst fractal dimension 0
label               0
dtype: int64
```

data\_frame.describe()

	mean radius	mean texture	mean perimeter	mean area	mean smoothness	mean compactness	mean concavity	mean concave points	mean symmetry	mean fractal dimension	...	
count	569.000000	569.000000	569.000000	569.000000	569.000000	569.000000	569.000000	569.000000	569.000000	569.000000	...	56
mean	14.127292	19.289649	91.969033	654.889104	0.096360	0.104341	0.088799	0.048919	0.181162	0.062798	...	2
std	3.524049	4.301036	24.298981	351.914129	0.014064	0.052813	0.079720	0.038803	0.027414	0.007060	...	
min	6.981000	9.710000	43.790000	143.500000	0.052630	0.019380	0.000000	0.000000	0.106000	0.049960	...	1
25%	11.700000	16.170000	75.170000	420.300000	0.086370	0.064920	0.029560	0.020310	0.161900	0.057700	...	2
50%	13.370000	18.840000	86.240000	551.100000	0.095870	0.092630	0.061540	0.033500	0.179200	0.061540	...	2
75%	15.780000	21.800000	104.100000	782.700000	0.105300	0.130400	0.130700	0.074000	0.195700	0.066120	...	2
max	28.110000	39.280000	188.500000	2501.000000	0.163400	0.345400	0.426800	0.201200	0.304000	0.097440	...	4

8 rows × 31 columns

data\_frame['label'].value\_counts()

```
1    357
0    212
Name: label, dtype: int64
```

data\_frame.groupby('label').mean()

	mean radius	mean texture	mean perimeter	mean area	mean smoothness	mean compactness	mean concavity	mean concave points	mean symmetry	mean fractal dimension	...	worst radius
label												
0	17.462830	21.604906	115.365377	978.376415	0.102898	0.145188	0.160775	0.087990	0.192909	0.062680	...	21.134811
1	12.146524	17.914762	78.075406	462.790196	0.092478	0.080085	0.046058	0.025717	0.174186	0.062867	...	13.379801

2 rows × 30 columns

```
X = data_frame.drop(columns='label', axis=1)
Y = data_frame['label']
```

print(X)

	mean radius	mean texture	mean perimeter	mean area	mean smoothness	\
0	17.99	10.38	122.80	1001.0	0.11840	
1	20.57	17.77	132.90	1326.0	0.08474	
2	19.69	21.25	130.00	1203.0	0.10960	
3	11.42	20.38	77.58	386.1	0.14250	
4	20.29	14.34	135.10	1297.0	0.10030	
..	...	...	...	...	...	
564	21.56	22.39	142.00	1479.0	0.11100	
565	20.13	28.25	131.20	1261.0	0.09780	
566	16.60	28.08	108.30	858.1	0.08455	
567	20.60	29.33	140.10	1265.0	0.11780	
568	7.76	24.54	47.92	181.0	0.05263	
	mean compactness	mean concavity	mean concave points	mean symmetry	\	
0	0.27760	0.30010	0.14710	0.2419		
1	0.07864	0.08690	0.07017	0.1812		
2	0.15990	0.19740	0.12790	0.2069		
3	0.28390	0.24140	0.10520	0.2597		
4	0.13280	0.19800	0.10430	0.1809		
..	...	...	...	...		

564	0.11590	0.24390	0.13890	0.1726
565	0.10340	0.14400	0.09791	0.1752
566	0.10230	0.09251	0.05302	0.1590
567	0.27700	0.35140	0.15200	0.2397
568	0.04362	0.00000	0.00000	0.1587

	mean fractal dimension	...	worst radius	worst texture \
0	0.07871	...	25.380	17.33
1	0.05667	...	24.990	23.41
2	0.05999	...	23.570	25.53
3	0.09744	...	14.910	26.50
4	0.05883	...	22.540	16.67
..	...	...	...	...
564	0.05623	...	25.450	26.40
565	0.05533	...	23.690	38.25
566	0.05648	...	18.980	34.12
567	0.07016	...	25.740	39.42
568	0.05884	...	9.456	30.37

	worst perimeter	worst area	worst smoothness	worst compactness \
0	184.60	2019.0	0.16220	0.66560
1	158.80	1956.0	0.12380	0.18660
2	152.50	1709.0	0.14440	0.42450
3	98.87	567.7	0.20980	0.86630
4	152.20	1575.0	0.13740	0.20500
..	...	...	...	...
564	166.10	2027.0	0.14100	0.21130
565	155.00	1731.0	0.11660	0.19220
566	126.70	1124.0	0.11390	0.30940
567	184.60	1821.0	0.16500	0.86810
568	59.16	268.6	0.08996	0.06444

	worst concavity	worst concave points	worst symmetry \
0	0.7119	0.2654	0.4601
1	0.2416	0.1860	0.2750
2	0.4504	0.2430	0.3613
3	0.6869	0.2575	0.6638
4	0.4000	0.1625	0.2364

```
print(Y)
```

```
0      0
1      0
2      0
3      0
4      0
..
564    0
565    0
566    0
567    0
568    1
Name: label, Length: 569, dtype: int64
```

```
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_state=2)
```

```
print(X.shape, X_train.shape, X_test.shape)
```

```
(569, 30) (455, 30) (114, 30)
```

```
from sklearn.preprocessing import StandardScaler
```

```
scaler = StandardScaler()
```

```
X_train_std = scaler.fit_transform(X_train)
```

```
X_test_std = scaler.transform(X_test)
```

```
import tensorflow as tf
tf.random.set_seed(3)
from tensorflow import keras
```

```
model = keras.Sequential([
    keras.layers.Flatten(input_shape=(30,)),
    keras.layers.Dense(20, activation='relu'),
    keras.layers.Dense(2, activation='sigmoid')
])
```

```
model.compile(optimizer='adam',
              loss='sparse_categorical_crossentropy',
              metrics=['accuracy'])
```

```
history = model.fit(X_train_std, Y_train, validation_split=0.1, epochs=10)
```

```
Epoch 1/10
13/13 [=====] - 1s 25ms/step - loss: 0.6659 - accuracy: 0.6822 - val_loss: 0.4986 - val_accuracy: 0.7826
Epoch 2/10
13/13 [=====] - 0s 7ms/step - loss: 0.4425 - accuracy: 0.7995 - val_loss: 0.3441 - val_accuracy: 0.8696
Epoch 3/10
13/13 [=====] - 0s 5ms/step - loss: 0.3243 - accuracy: 0.8826 - val_loss: 0.2706 - val_accuracy: 0.8913
Epoch 4/10
13/13 [=====] - 0s 6ms/step - loss: 0.2650 - accuracy: 0.9120 - val_loss: 0.2284 - val_accuracy: 0.9130
Epoch 5/10
13/13 [=====] - 0s 6ms/step - loss: 0.2281 - accuracy: 0.9267 - val_loss: 0.2001 - val_accuracy: 0.9130
Epoch 6/10
13/13 [=====] - 0s 6ms/step - loss: 0.2010 - accuracy: 0.9315 - val_loss: 0.1795 - val_accuracy: 0.9565
Epoch 7/10
13/13 [=====] - 0s 5ms/step - loss: 0.1793 - accuracy: 0.9413 - val_loss: 0.1632 - val_accuracy: 0.9565
Epoch 8/10
13/13 [=====] - 0s 5ms/step - loss: 0.1627 - accuracy: 0.9438 - val_loss: 0.1493 - val_accuracy: 0.9565
Epoch 9/10
13/13 [=====] - 0s 4ms/step - loss: 0.1485 - accuracy: 0.9511 - val_loss: 0.1375 - val_accuracy: 0.9565
Epoch 10/10
13/13 [=====] - 0s 6ms/step - loss: 0.1362 - accuracy: 0.9560 - val_loss: 0.1292 - val_accuracy: 0.9565
```

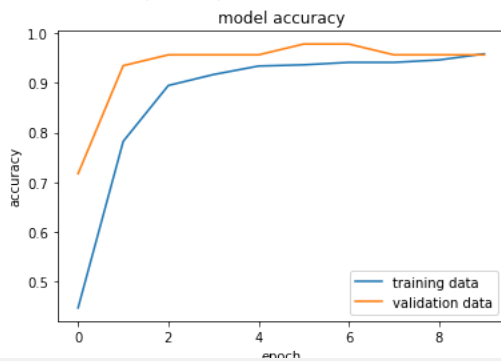
## Visualizing accuracy and loss

```
plt.plot(history.history['accuracy'])
plt.plot(history.history['val_accuracy'])

plt.title('model accuracy')
plt.ylabel('accuracy')
plt.xlabel('epoch')

plt.legend(['training data', 'validation data'], loc = 'lower right')
```

```
<matplotlib.legend.Legend at 0x7fcf3ff52450>
```

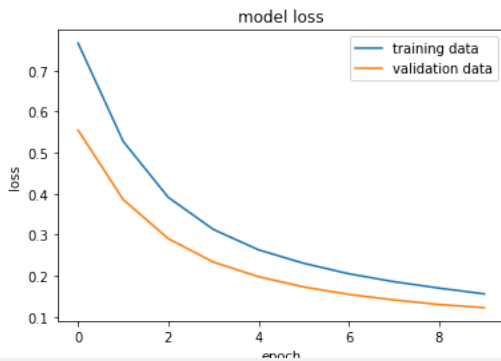


```
plt.plot(history.history['loss'])
plt.plot(history.history['val_loss'])

plt.title('model loss')
plt.ylabel('loss')
plt.xlabel('epoch')

plt.legend(['training data', 'validation data'], loc = 'upper right')
```

```
<matplotlib.legend.Legend at 0x7fcf409fbcd0>
```



## Accuracy of the model on test data

```
loss, accuracy = model.evaluate(X_test_std, Y_test)
print(accuracy)

4/4 [=====] - 0s 5ms/step - loss: 0.1693 - accuracy: 0.9386
0.9385964870452881
```

```
print(X_test_std.shape)
print(X_test_std[0])

(114, 30)
[-0.04462793 -1.41612656 -0.05903514 -0.16234067  2.0202457 -0.11323672
 0.18500609  0.47102419  0.63336386  0.26335737  0.53209124  2.62763999
 0.62351167  0.11405261  1.01246781  0.41126289  0.63848593  2.88971815
 -0.41675911  0.74270853 -0.32983699 -1.67435595 -0.36854552 -0.38767294
 0.32655007 -0.74858917 -0.54689089 -0.18278004 -1.23064515 -0.6268286 ]
```

```
Y_pred = model.predict(X_test_std)
```

```
print(Y_pred.shape)
print(Y_pred[0])

(114, 2)
[0.24822891 0.537705 ]
```

```
print(X_test_std)

[[-0.04462793 -1.41612656 -0.05903514 ... -0.18278004 -1.23064515
  -0.6268286 ]
 [ 0.24583601 -0.06219797  0.21802678 ...  0.54129749  0.11047691
  0.0483572 ]
 [-1.26115925 -0.29051645 -1.26499659 ... -1.35138617  0.269338
  -0.28231213]
 ...
 [ 0.72709489  0.45836817  0.75277276 ...  1.46701686  1.19909344
  0.65319961]
 [ 0.25437907  1.33054477  0.15659489 ... -1.29043534 -2.22561725
 -1.59557344]
 [ 0.84100232 -0.06676434  0.8929529 ...  2.15137705  0.35629355
  0.37459546]]
```

```
print(Y_pred)

[[2.48228908e-01 5.37705004e-01]
 [5.04756272e-01 5.95336974e-01]
 [3.67398351e-01 9.86996770e-01]
 [9.29620504e-01 2.29299068e-04]
 [5.93764186e-01 4.93302941e-01]
 [8.53436947e-01 1.02660358e-02]
 [2.81091124e-01 6.12119973e-01]
 [4.79383349e-01 9.91485417e-01]
 [4.34603453e-01 9.51259434e-01]
 [6.34022474e-01 9.53936636e-01]
 [4.83262986e-01 6.63963675e-01]
 [3.86232853e-01 9.23828781e-01]
 [2.77723640e-01 7.82129407e-01]
 [3.78485560e-01 8.21216226e-01]
 [4.91592646e-01 9.53726530e-01]
 [7.36154735e-01 1.57864958e-01]
 [4.71651524e-01 9.74555612e-01]
 [2.54375935e-01 8.85443091e-01]
 [5.75661778e-01 9.49565291e-01]
 [8.55701387e-01 1.94370747e-02]
 [1.90429598e-01 1.22809112e-02]
 [4.47704107e-01 9.82434034e-01]
 [5.11266708e-01 9.58328366e-01]
 [5.38256645e-01 9.74644005e-01]
 [5.29228508e-01 8.16611886e-01]
 [7.13969886e-01 6.25776052e-02]
 [4.54092264e-01 9.09015238e-01]
 [5.10061681e-01 7.40805507e-01]
 [6.70265079e-01 1.68987542e-01]
 [6.56581044e-01 1.18384689e-01]
 [5.91236949e-01 9.38910842e-01]
 [4.20632422e-01 9.19608235e-01]
 [5.11391044e-01 9.51588750e-01]
 [8.78562212e-01 1.91476941e-03]
 [8.40861440e-01 3.40589881e-02]
 [5.59490383e-01 7.92683899e-01]
 [4.08323050e-01 9.93038535e-01]
 [5.17861009e-01 8.50202918e-01]
 [3.98939788e-01 9.78648603e-01]
 [3.16001832e-01 9.14353132e-01]
 [9.31815386e-01 1.04099512e-03]
 [5.36973536e-01 2.56363392e-01]
 [3.96728277e-01 9.79030609e-01]]
```

```
[2.11794227e-01 8.85924697e-01]
[7.01928675e-01 1.99133068e-01]
[4.22424793e-01 9.64499712e-01]
[2.93639600e-01 9.86928105e-01]
[4.57987458e-01 9.43145216e-01]
[8.21546316e-01 1.38805807e-02]
[7.11578608e-01 1.09298050e-01]
[6.50350690e-01 9.58284259e-01]
[6.42665863e-01 3.17812234e-01]
[5.54926872e-01 6.52379811e-01]
[4.63142991e-01 9.63462353e-01]
[3.21405470e-01 9.79116678e-01]
[5.65109372e-01 6.47464037e-01]
[4.16883409e-01 8.96000743e-01]
[2.26916492e-01 9.79369640e-01]
```

model.predict() gives the prediction probability of each class for that data point

```
# argmax function
```

```
my_list = [0.25, 0.56]
```

```
index_of_max_value = np.argmax(my_list)
print(my_list)
print(index_of_max_value)
```

```
→ [0.25, 0.56]
1
```

```
# converting the prediction probability to class labels
```

```
Y_pred_labels = [np.argmax(i) for i in Y_pred]
print(Y_pred_labels)
```

```
→ [1, 1, 1, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 1, 1, 0, 0, 1, 1,
```

## Building the predictive system

```
input_data = (11.76,21.6,74.72,427.9,0.08637,0.04966,0.01657,0.01115,0.1495,0.05888,0.4062,1.21,2.635,28.47,0.005857,0.009758,0.01168,0
```

```
# change the input_data to a numpy array
input_data_as_numpy_array = np.asarray(input_data)
```

```
# reshape the numpy array as we are predicting for one data point
input_data_reshaped = input_data_as_numpy_array.reshape(1,-1)
```

```
# standardizing the input data
input_data_std = scaler.transform(input_data_reshaped)
```

```
prediction = model.predict(input_data_std)
print(prediction)
```

```
prediction_label = [np.argmax(prediction)]
print(prediction_label)
```

```
if(prediction_label[0] == 0):
    print('The tumor is Malignant')
```

```
else:
    print('The tumor is Benign')
```

```
→ [[0.42537233 0.9805447 ]]
[1]
```

```
The tumor is Benign
```

```
/usr/local/lib/python3.7/dist-packages/sklearn/base.py:451: UserWarning: X does not have valid feature names, but StandardScaler was
"X does not have valid feature names, but"
```

Start coding or [generate](#) with AI.

