

Disease Predictor using Data Mining

Dataset link:

<https://www.kaggle.com/itachi9604/disease-symptom-description-dataset?select=dataset.csv>

Tools and Techniques

- Environment used is PyCharm
- Classifier Algorithm used is Support Vector Classifier
- Frontend framework is PyQt5
- Backend Language used is Python 3.6
- Used the pandas framework for data manipulation and cleaning

Importing libraries

```
In [1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.svm import SVC
from sklearn.metrics import f1_score, accuracy_score, confusion_matrix
import seaborn as sns
```

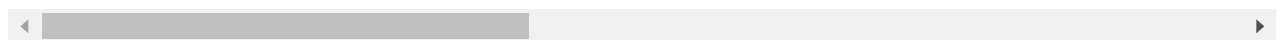
Read the dataset

```
In [2]: df = pd.read_csv('data/disease_dataset.csv')
df.head(10)
```

```
Out[2]:
```

	Disease	Symptom_1	Symptom_2	Symptom_3	Symptom_4	Symptom_5	Symptom_6
0	Fungal infection	itching	skin_rash	nodal_skin_eruptions	dischromic _patches	NaN	NaN
1	Fungal infection	skin_rash	nodal_skin_eruptions	dischromic _patches	NaN	NaN	NaN
2	Fungal infection	itching	nodal_skin_eruptions	dischromic _patches	NaN	NaN	NaN
3	Fungal infection	itching	skin_rash	dischromic _patches	NaN	NaN	NaN
4	Fungal infection	itching	skin_rash	nodal_skin_eruptions	NaN	NaN	NaN
5	Fungal infection	skin_rash	nodal_skin_eruptions	dischromic _patches	NaN	NaN	NaN
6	Fungal infection	itching	nodal_skin_eruptions	dischromic _patches	NaN	NaN	NaN
7	Fungal infection	itching	skin_rash	dischromic _patches	NaN	NaN	NaN

	Disease	Symptom_1	Symptom_2	Symptom_3	Symptom_4	Symptom_5	Symptom_6
8	Fungal infection	itching	skin_rash	nodal_skin_eruptions	NaN	NaN	NaN
9	Fungal infection	itching	skin_rash	nodal_skin_eruptions	dischromic _patches	NaN	NaN



Dataset info

In [3]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4923 entries, 0 to 4922
Data columns (total 18 columns):
#   Column          Non-Null Count  Dtype
---  -
0   Disease         4923 non-null   object
1   Symptom_1       4923 non-null   object
2   Symptom_2       4923 non-null   object
3   Symptom_3       4923 non-null   object
4   Symptom_4       4575 non-null   object
5   Symptom_5       3716 non-null   object
6   Symptom_6       2936 non-null   object
7   Symptom_7       2270 non-null   object
8   Symptom_8       1946 non-null   object
9   Symptom_9       1693 non-null   object
10  Symptom_10      1512 non-null   object
11  Symptom_11      1194 non-null   object
12  Symptom_12      744 non-null    object
13  Symptom_13      504 non-null    object
14  Symptom_14      306 non-null    object
15  Symptom_15      240 non-null    object
16  Symptom_16      192 non-null    object
17  Symptom_17      72 non-null     object
dtypes: object(18)
memory usage: 692.4+ KB
```

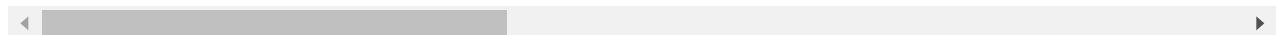
Dataset characteristics

In [4]:

```
df.describe()
```

Out[4]:

	Disease	Symptom_1	Symptom_2	Symptom_3	Symptom_4	Symptom_5	Symptom_6
count	4923	4923	4923	4923	4575	3716	2936
unique	42	36	50	56	52	40	34
top	Hypertension	vomiting	vomiting	fatigue	high_fever	headache	nausea abd
freq	120	822	870	726	378	348	390



Check for null and NaN values

In [5]:

```
df.isna().sum()
```

Out[5]: Disease

0

```
Symptom_1      0
Symptom_2      0
Symptom_3      0
Symptom_4     348
Symptom_5    1207
Symptom_6    1987
Symptom_7    2653
Symptom_8    2977
Symptom_9    3230
Symptom_10    3411
Symptom_11    3729
Symptom_12    4179
Symptom_13    4419
Symptom_14    4617
Symptom_15    4683
Symptom_16    4731
Symptom_17    4851
dtype: int64
```

Remove the trailing space from the symptom columns

```
In [6]: cols = df.columns
data = df[cols].values.flatten()

s = pd.Series(data)
s = s.str.strip()
s = s.values.reshape(df.shape)

df = pd.DataFrame(s, columns=df.columns)
df.head()
```

```
Out[6]:
```

	Disease	Symptom_1	Symptom_2	Symptom_3	Symptom_4	Symptom_5	Symptom_6
0	Fungal infection	itching	skin_rash	nodal_skin_eruptions	dischromic _patches	NaN	NaN
1	Fungal infection	skin_rash	nodal_skin_eruptions	dischromic _patches	NaN	NaN	NaN
2	Fungal infection	itching	nodal_skin_eruptions	dischromic _patches	NaN	NaN	NaN
3	Fungal infection	itching	skin_rash	dischromic _patches	NaN	NaN	NaN
4	Fungal infection	itching	skin_rash	nodal_skin_eruptions	NaN	NaN	NaN

Total Disease are

```
In [7]: df['Disease'].unique()

Out[7]: array(['Fungal infection', 'Allergy', 'GERD', 'Chronic cholestasis',
        'Drug Reaction', 'Peptic ulcer diseae', 'AIDS', 'Diabetes',
        'Gastroenteritis', 'Bronchial Asthma', 'Hypertension', 'Migraine',
        'Cervical spondylosis', 'Paralysis (brain hemorrhage)', 'Jaundice',
        'Malaria', 'Chicken pox', 'Dengue', 'Typhoid', 'hepatitis A',
        'Hepatitis B', 'Hepatitis C', 'Hepatitis D', 'Hepatitis E',
        'Alcoholic hepatitis', 'Tuberculosis', 'Common Cold', 'Pneumonia',
        'Dimorphic hemmorhoids(piles)', 'Heart attack', 'Varicose veins',
```

```
'Hypothyroidism', 'Hyperthyroidism', 'Hypoglycemia',
'Osteoarthritis', 'Arthritis',
'(vertigo) Parosymal Positional Vertigo', 'Acne',
'Urinary tract infection', 'Psoriasis', 'Impetigo', 'covid'],
dtype=object)
```

Fill the NaN values with zero

```
In [8]: df = df.fillna(0)
df.head()
```

```
Out[8]:
```

	Disease	Symptom_1	Symptom_2	Symptom_3	Symptom_4	Symptom_5	Symptom_6
0	Fungal infection	itching	skin_rash	nodal_skin_eruptions	dischromic _patches	0	0
1	Fungal infection	skin_rash	nodal_skin_eruptions	dischromic _patches	0	0	0
2	Fungal infection	itching	nodal_skin_eruptions	dischromic _patches	0	0	0
3	Fungal infection	itching	skin_rash	dischromic _patches	0	0	0
4	Fungal infection	itching	skin_rash	nodal_skin_eruptions	0	0	0

Encode symptoms in the data with the symptom rank

```
In [9]: from sklearn.preprocessing import LabelEncoder
le = LabelEncoder()
for column in df.columns:
    df[column] = le.fit_transform(df[column].astype(str))
```

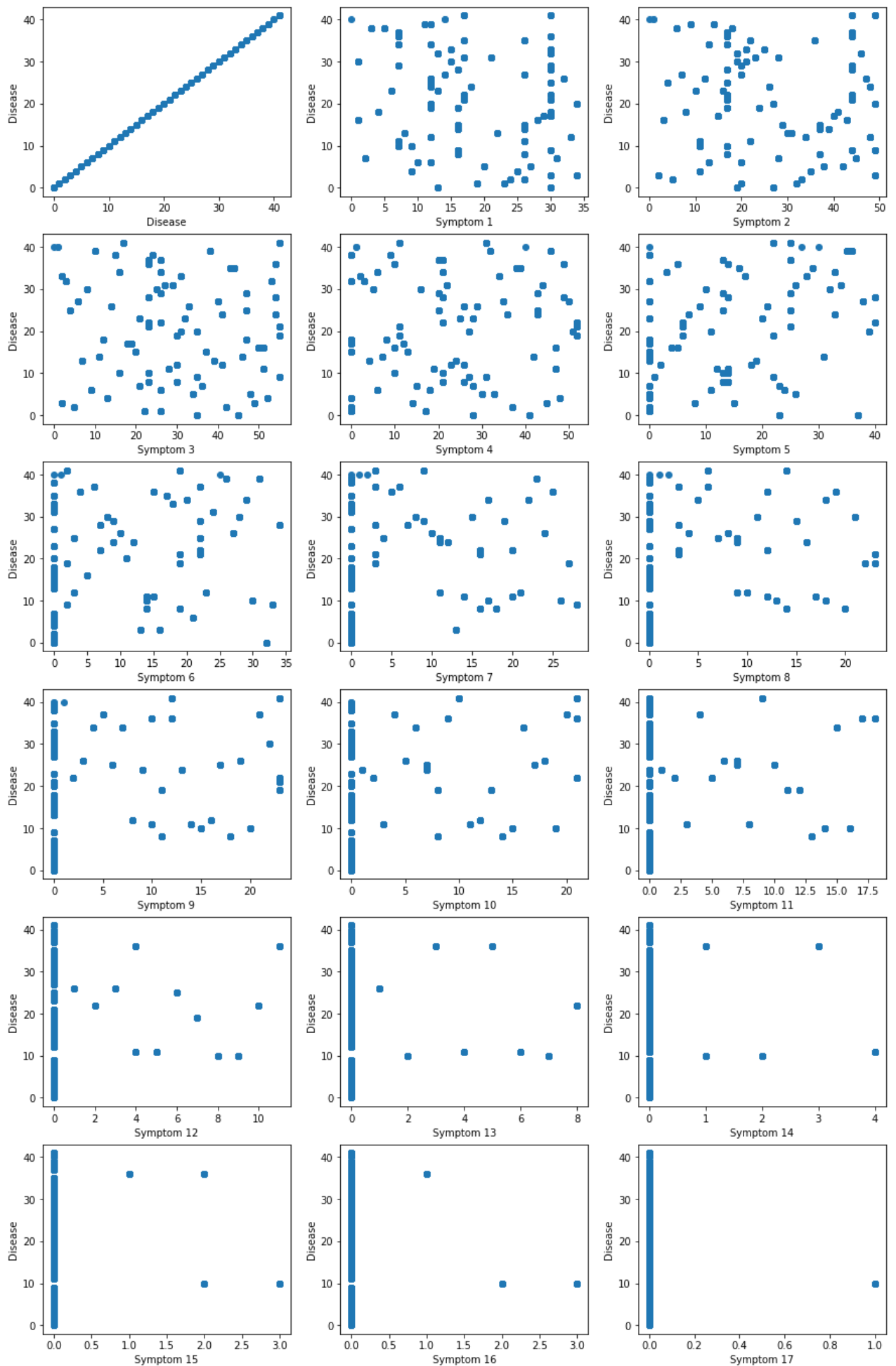
Plot the dataset attribute with each other

```
In [10]: import matplotlib.pyplot as plt

plt.figure(figsize=(15,45))

plot_count = 1
for feature in list(df.columns):
    plt.subplot(11,3,plot_count)
    plt.scatter(df[feature], df['Disease'])
    plt.xlabel(feature.replace('_', ' ').title())
    plt.ylabel('Disease')
    plot_count+=1

plt.show()
```



Check if entire columns have zero values so we can drop those values

```
In [11]: (df[cols] == 0).all()
```

```
Out[11]: Disease      False
Symptom_1      False
Symptom_2      False
Symptom_3      False
Symptom_4      False
Symptom_5      False
Symptom_6      False
Symptom_7      False
Symptom_8      False
Symptom_9      False
Symptom_10     False
Symptom_11     False
Symptom_12     False
Symptom_13     False
Symptom_14     False
Symptom_15     False
Symptom_16     False
Symptom_17     False
dtype: bool
```

Select the features as symptoms column and label as Disease column

```
In [12]: data = df.iloc[:,1:].values
labels = df['Disease'].values
```

Split the data, labels with shuffle

```
In [13]: x_train, x_test, y_train, y_test = train_test_split(data, labels, shuffle=True, train_s
print(x_train.shape, x_test.shape, y_train.shape, y_test.shape)
```

```
(4184, 17) (739, 17) (4184,) (739,)
```

Initialize and train a Support vector classifier

```
In [14]: model = SVC()
model.fit(x_train, y_train)
```

```
Out[14]: SVC()
```

Throw predictions

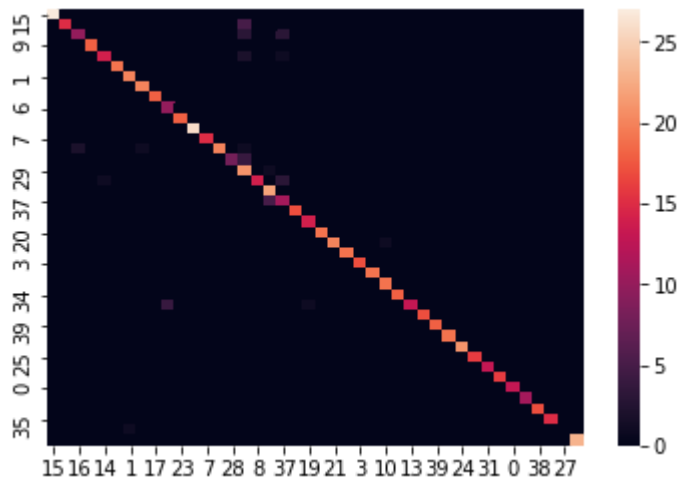
```
In [15]: preds = model.predict(x_test)
```

Plot the confusion matrix for 25 diseases and calculate f1, accuracy

```
In [16]: conf_mat = confusion_matrix(y_test, preds)
df_cm = pd.DataFrame(conf_mat, index=df['Disease'].unique(), columns=df['Disease'].unique())
print('F1-score% =', f1_score(y_test, preds, average='macro')*100, '|', 'Accuracy% =',
sns.heatmap(df_cm))
```

```
F1-score% = 92.42967792825483 | Accuracy% = 94.72259810554804
```

Out[16]: <AxesSubplot:>

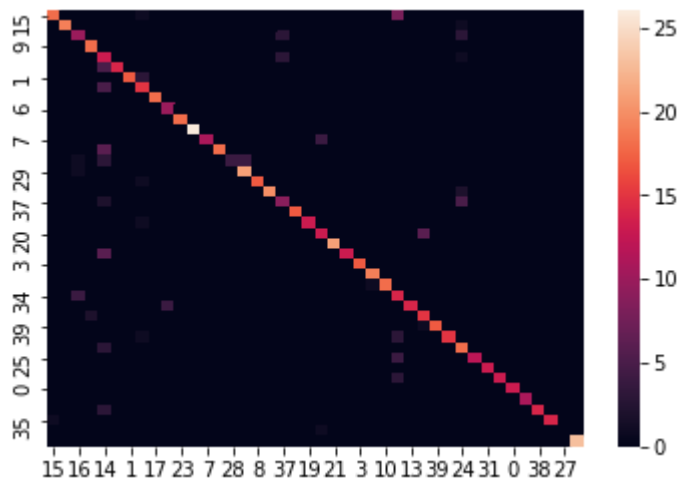


Now by using Naive Bayes classifier

```
In [19]: from sklearn.naive_bayes import GaussianNB
model = GaussianNB()
model.fit(x_train, y_train)
preds = model.predict(x_test)
conf_mat = confusion_matrix(y_test, preds)
df_cm = pd.DataFrame(conf_mat, index=df['Disease'].unique(), columns=df['Disease'].unique())
print('F1-score% =', f1_score(y_test, preds, average='macro')*100, '|', 'Accuracy% =',
      sns.heatmap(df_cm))
```

F1-score% = 84.45298953653445 | Accuracy% = 85.65629228687416

Out[19]: <AxesSubplot:>



Testing

Reading dataset of disease

```
In [1]: # read the dataset
import pandas as pd
import numpy as np
# Let assume the patient have the following symptoms
df = pd.read_csv("data/disease_dataset.csv")
```

remove all the spaces

```
In [2]: cols = df.columns
data = df[cols].values.flatten()

s = pd.Series(data)
s = s.str.strip()
s = s.values.reshape(df.shape)

df = pd.DataFrame(s, columns=df.columns)
```

Input and compare

```
In [3]: # Here is the actual implementation
```

	Disease	Symptom_1 \
0	Fungal infection	itching
1	Fungal infection	skin_rash
2	Fungal infection	itching
3	Fungal infection	itching
4	Fungal infection	itching
5	Fungal infection	skin_rash
6	Fungal infection	itching
7	Fungal infection	itching
8	Fungal infection	itching
9	Fungal infection	itching
10	Allergy	continuous_sneezing
11	Allergy	shivering
12	Allergy	continuous_sneezing
13	Allergy	continuous_sneezing
14	Allergy	continuous_sneezing
15	Allergy	shivering
16	Allergy	continuous_sneezing
17	Allergy	continuous_sneezing
18	Allergy	continuous_sneezing
19	Allergy	continuous_sneezing
20	GERD	stomach_pain
21	GERD	stomach_pain
22	GERD	stomach_pain
23	GERD	stomach_pain
24	GERD	stomach_pain
25	GERD	stomach_pain
26	GERD	acidity
27	GERD	stomach_pain
28	GERD	stomach_pain
29	GERD	stomach_pain
...
4893	Jaundice	itching
4894	Malaria	chills
4895	Chicken pox	itching
4896	Dengue	skin_rash
4897	Typhoid	chills
4898	hepatitis A	joint_pain
4899	Hepatitis B	itching
4900	Hepatitis C	fatigue
4901	Hepatitis D	joint_pain
4902	Hepatitis E	joint_pain
4903	Alcoholic hepatitis	vomiting
4904	Tuberculosis	chills
4905	Common Cold	continuous_sneezing

4906		Pneumonia	chills
4907	Dimorphic hemmorrhoids(piles)		constipation
4908		Heart attack	vomiting
4909		Varicose veins	fatigue
4910		Hypothyroidism	fatigue
4911		Hyperthyroidism	fatigue
4912		Hypoglycemia	vomiting
4913		Osteoarthritis	joint_pain
4914		Arthritis	muscle_weakness
4915	(vertigo) Paroymsal	Positional Vertigo	vomiting
4916		Acne	skin_rash
4917	Urinary tract	infection	burning_micturition
4918		Psoriasis	skin_rash
4919		Impetigo	skin_rash
4920		covid	high_fever
4921		covid	high_fever
4922		covid	Tiredness

	Symptom_2	Symptom_3 \
0	skin_rash	nodal_skin_eruptions
1	nodal_skin_eruptions	dischromic _patches
2	nodal_skin_eruptions	dischromic _patches
3	skin_rash	dischromic _patches
4	skin_rash	nodal_skin_eruptions
5	nodal_skin_eruptions	dischromic _patches
6	nodal_skin_eruptions	dischromic _patches
7	skin_rash	dischromic _patches
8	skin_rash	nodal_skin_eruptions
9	skin_rash	nodal_skin_eruptions
10	shivering	chills
11	chills	watering_from_eyes
12	chills	watering_from_eyes
13	shivering	watering_from_eyes
14	shivering	chills
15	chills	watering_from_eyes
16	chills	watering_from_eyes
17	shivering	watering_from_eyes
18	shivering	chills
19	shivering	chills
20	acidity	ulcers_on_tongue
21	ulcers_on_tongue	vomiting
22	acidity	vomiting
23	acidity	ulcers_on_tongue
24	acidity	ulcers_on_tongue
25	acidity	ulcers_on_tongue
26	ulcers_on_tongue	vomiting
27	ulcers_on_tongue	vomiting
28	acidity	vomiting
29	acidity	ulcers_on_tongue
...
4893	vomiting	fatigue
4894	vomiting	high_fever
4895	skin_rash	fatigue
4896	chills	joint_pain
4897	vomiting	fatigue
4898	vomiting	yellowish_skin
4899	fatigue	lethargy
4900	yellowish_skin	nausea
4901	vomiting	fatigue
4902	vomiting	fatigue
4903	yellowish_skin	abdominal_pain
4904	vomiting	fatigue
4905	chills	fatigue
4906	fatigue	cough
4907	pain_during_bowel_movements	pain_in_anal_region

4908	breathlessness	sweating
4909	cramps	bruising
4910	weight_gain	cold_hands_and_feets
4911	mood_swings	weight_loss
4912	fatigue	anxiety
4913	neck_pain	knee_pain
4914	stiff_neck	swelling_joints
4915	headache	nausea
4916	pus_filled_pimples	blackheads
4917	bladder_discomfort	foul_smell_of urine
4918	joint_pain	skin_peeling
4919	high_fever	blister
4920	Tiredness	Dry-Cough
4921	Tiredness	Dry-Cough
4922	Dry-Cough	Difficulty-in-Breathing

	Symptom_4	Symptom_5 \
0	dischromic _patches	NaN
1	NaN	NaN
2	NaN	NaN
3	NaN	NaN
4	NaN	NaN
5	NaN	NaN
6	NaN	NaN
7	NaN	NaN
8	NaN	NaN
9	dischromic _patches	NaN
10	watering_from_eyes	NaN
11	NaN	NaN
12	NaN	NaN
13	NaN	NaN
14	NaN	NaN
15	NaN	NaN
16	NaN	NaN
17	NaN	NaN
18	NaN	NaN
19	watering_from_eyes	NaN
20	vomiting	cough
21	cough	chest_pain
22	cough	chest_pain
23	cough	chest_pain
24	vomiting	chest_pain
25	vomiting	cough
26	cough	chest_pain
27	cough	chest_pain
28	cough	chest_pain
29	cough	chest_pain
...
4893	weight_loss	high_fever
4894	sweating	headache
4895	lethargy	high_fever
4896	vomiting	fatigue
4897	high_fever	headache
4898	dark_urine	nausea
4899	yellowish_skin	dark_urine
4900	loss_of_appetite	yellowing_of_eyes
4901	yellowish_skin	dark_urine
4902	high_fever	yellowish_skin
4903	swelling_of_stomach	distention_of_abdomen
4904	weight_loss	cough
4905	cough	high_fever
4906	high_fever	breathlessness
4907	bloody_stool	irritation_in_anus
4908	chest_pain	NaN
4909	obesity	swollen_legs

4910	mood_swings	lethargy
4911	restlessness	sweating
4912	sweating	headache
4913	hip_joint_pain	swelling_joints
4914	movement_stiffness	painful_walking
4915	spinning_movements	loss_of_balance
4916	scurrying	NaN
4917	continuous_feel_of_urine	NaN
4918	silver_like_dusting	small_dents_in_nails
4919	red_sore_around_nose	yellow_crust_ooze
4920	Difficulty-in-Breathing	sore_throat
4921	Difficulty-in-Breathing	NaN
4922	sore_throat	pains

	Symptom_6	Symptom_7 \
0	NaN	NaN
1	NaN	NaN
2	NaN	NaN
3	NaN	NaN
4	NaN	NaN
5	NaN	NaN
6	NaN	NaN
7	NaN	NaN
8	NaN	NaN
9	NaN	NaN
10	NaN	NaN
11	NaN	NaN
12	NaN	NaN
13	NaN	NaN
14	NaN	NaN
15	NaN	NaN
16	NaN	NaN
17	NaN	NaN
18	NaN	NaN
19	NaN	NaN
20	chest_pain	NaN
21	NaN	NaN
22	NaN	NaN
23	NaN	NaN
24	NaN	NaN
25	NaN	NaN
26	NaN	NaN
27	NaN	NaN
28	NaN	NaN
29	NaN	NaN
...
4893	yellowish_skin	dark_urine
4894	nausea	diarrhoea
4895	headache	loss_of_appetite
4896	high_fever	headache
4897	nausea	constipation
4898	loss_of_appetite	abdominal_pain
4899	loss_of_appetite	abdominal_pain
4900	family_history	NaN
4901	nausea	loss_of_appetite
4902	dark_urine	nausea
4903	history_of_alcohol_consumption	fluid_overload
4904	high_fever	breathlessness
4905	headache	swelled_lymph_nodes
4906	sweating	malaise
4907	NaN	NaN
4908	NaN	NaN
4909	swollen_blood_vessels	prominent_veins_on_calf
4910	dizziness	puffy_face_and_eyes
4911	diarrhoea	fast_heart_rate

4912		nausea	blurred_and_distorted_vision	
4913		painful_walking		NaN
4914		NaN		NaN
4915		unsteadiness		NaN
4916		NaN		NaN
4917		NaN		NaN
4918		inflammatory_nails		NaN
4919		NaN		NaN
4920		pains	Nasal_Congestion	
4921		NaN		NaN
4922		Nasal_Congestion		Runny_Nose

	Symptom_8	Symptom_9	Symptom_10	\
0	NaN	NaN	NaN	
1	NaN	NaN	NaN	
2	NaN	NaN	NaN	
3	NaN	NaN	NaN	
4	NaN	NaN	NaN	
5	NaN	NaN	NaN	
6	NaN	NaN	NaN	
7	NaN	NaN	NaN	
8	NaN	NaN	NaN	
9	NaN	NaN	NaN	
10	NaN	NaN	NaN	
11	NaN	NaN	NaN	
12	NaN	NaN	NaN	
13	NaN	NaN	NaN	
14	NaN	NaN	NaN	
15	NaN	NaN	NaN	
16	NaN	NaN	NaN	
17	NaN	NaN	NaN	
18	NaN	NaN	NaN	
19	NaN	NaN	NaN	
20	NaN	NaN	NaN	
21	NaN	NaN	NaN	
22	NaN	NaN	NaN	
23	NaN	NaN	NaN	
24	NaN	NaN	NaN	
25	NaN	NaN	NaN	
26	NaN	NaN	NaN	
27	NaN	NaN	NaN	
28	NaN	NaN	NaN	
29	NaN	NaN	NaN	
...	
4893	abdominal_pain	NaN	NaN	
4894	muscle_pain	NaN	NaN	
4895	mild_fever	swelled_lymph_nodes	malaise	
4896	nausea	loss_of_appetite	pain_behind_the_eyes	
4897	abdominal_pain	diarrhoea	toxic_look_(typhos)	
4898	diarrhoea	mild_fever	yellowing_of_eyes	
4899	yellow_urine	yellowing_of_eyes	malaise	
4900	NaN	NaN	NaN	
4901	abdominal_pain	yellowing_of_eyes	NaN	
4902	loss_of_appetite	abdominal_pain	yellowing_of_eyes	
4903	NaN	NaN	NaN	
4904	sweating	loss_of_appetite	mild_fever	
4905	malaise	phlegm	throat_irritation	
4906	phlegm	chest_pain	fast_heart_rate	
4907	NaN	NaN	NaN	
4908	NaN	NaN	NaN	
4909	NaN	NaN	NaN	
4910	enlarged_thyroid	brittle_nails	swollen_extremeties	
4911	excessive_hunger	muscle_weakness	irritability	
4912	excessive_hunger	drying_and_tingling_lips	slurred_speech	
4913	NaN	NaN	NaN	

4914	NaN	NaN	NaN
4915	NaN	NaN	NaN
4916	NaN	NaN	NaN
4917	NaN	NaN	NaN
4918	NaN	NaN	NaN
4919	NaN	NaN	NaN
4920	Runny_Nose	Diarrhea	NaN
4921	NaN	NaN	NaN
4922	Diarrhea	NaN	NaN

	Symptom_11	Symptom_12 \
0	NaN	NaN
1	NaN	NaN
2	NaN	NaN
3	NaN	NaN
4	NaN	NaN
5	NaN	NaN
6	NaN	NaN
7	NaN	NaN
8	NaN	NaN
9	NaN	NaN
10	NaN	NaN
11	NaN	NaN
12	NaN	NaN
13	NaN	NaN
14	NaN	NaN
15	NaN	NaN
16	NaN	NaN
17	NaN	NaN
18	NaN	NaN
19	NaN	NaN
20	NaN	NaN
21	NaN	NaN
22	NaN	NaN
23	NaN	NaN
24	NaN	NaN
25	NaN	NaN
26	NaN	NaN
27	NaN	NaN
28	NaN	NaN
29	NaN	NaN
...
4893	NaN	NaN
4894	NaN	NaN
4895	red_spots_over_body	NaN
4896	back_pain	malaise
4897	belly_pain	NaN
4898	muscle_pain	NaN
4899	receiving_blood_transfusion	receiving_unsterile_injections
4900	NaN	NaN
4901	NaN	NaN
4902	acute_liver_failure	coma
4903	NaN	NaN
4904	yellowing_of_eyes	swelled_lymph_nodes
4905	redness_of_eyes	sinus_pressure
4906	rusty_sputum	NaN
4907	NaN	NaN
4908	NaN	NaN
4909	NaN	NaN
4910	depression	irritability
4911	abnormal_menstruation	NaN
4912	irritability	palpitations
4913	NaN	NaN
4914	NaN	NaN
4915	NaN	NaN

4916	NaN	NaN
4917	NaN	NaN
4918	NaN	NaN
4919	NaN	NaN
4920	NaN	NaN
4921	NaN	NaN
4922	NaN	NaN

	Symptom_13	Symptom_14	Symptom_15	Symptom_16 \
0	NaN	NaN	NaN	NaN
1	NaN	NaN	NaN	NaN
2	NaN	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN
5	NaN	NaN	NaN	NaN
6	NaN	NaN	NaN	NaN
7	NaN	NaN	NaN	NaN
8	NaN	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN
17	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN
23	NaN	NaN	NaN	NaN
24	NaN	NaN	NaN	NaN
25	NaN	NaN	NaN	NaN
26	NaN	NaN	NaN	NaN
27	NaN	NaN	NaN	NaN
28	NaN	NaN	NaN	NaN
29	NaN	NaN	NaN	NaN
...
4893	NaN	NaN	NaN	NaN
4894	NaN	NaN	NaN	NaN
4895	NaN	NaN	NaN	NaN
4896	muscle_pain	red_spots_over_body	NaN	NaN
4897	NaN	NaN	NaN	NaN
4898	NaN	NaN	NaN	NaN
4899	NaN	NaN	NaN	NaN
4900	NaN	NaN	NaN	NaN
4901	NaN	NaN	NaN	NaN
4902	stomach_bleeding	NaN	NaN	NaN
4903	NaN	NaN	NaN	NaN
4904	malaise	phlegm	chest_pain	blood_in_sputum
4905	runny_nose	congestion	chest_pain	loss_of_smell
4906	NaN	NaN	NaN	NaN
4907	NaN	NaN	NaN	NaN
4908	NaN	NaN	NaN	NaN
4909	NaN	NaN	NaN	NaN
4910	abnormal_menstruation	NaN	NaN	NaN
4911	NaN	NaN	NaN	NaN
4912	NaN	NaN	NaN	NaN
4913	NaN	NaN	NaN	NaN
4914	NaN	NaN	NaN	NaN
4915	NaN	NaN	NaN	NaN
4916	NaN	NaN	NaN	NaN
4917	NaN	NaN	NaN	NaN

4918	NaN	NaN	NaN	NaN
4919	NaN	NaN	NaN	NaN
4920	NaN	NaN	NaN	NaN
4921	NaN	NaN	NaN	NaN
4922	NaN	NaN	NaN	NaN

	Symptom_17
--	------------

0	NaN
1	NaN
2	NaN
3	NaN
4	NaN
5	NaN
6	NaN
7	NaN
8	NaN
9	NaN
10	NaN
11	NaN
12	NaN
13	NaN
14	NaN
15	NaN
16	NaN
17	NaN
18	NaN
19	NaN
20	NaN
21	NaN
22	NaN
23	NaN
24	NaN
25	NaN
26	NaN
27	NaN
28	NaN
29	NaN
...	...
4893	NaN
4894	NaN
4895	NaN
4896	NaN
4897	NaN
4898	NaN
4899	NaN
4900	NaN
4901	NaN
4902	NaN
4903	NaN
4904	NaN
4905	muscle_pain
4906	NaN
4907	NaN
4908	NaN
4909	NaN
4910	NaN
4911	NaN
4912	NaN
4913	NaN
4914	NaN
4915	NaN
4916	NaN
4917	NaN
4918	NaN
4919	NaN

```
4920      NaN
4921      NaN
4922      NaN
```

```
[4923 rows x 18 columns]
```

```
      Disease Symptom_1 Symptom_2 Symptom_3 Symptom_4 \
4922   covid  Tiredness Dry-Cough Difficulty-in-Breathing sore_throat

      Symptom_5 Symptom_6 Symptom_7 Symptom_8 Symptom_9 Symptom_10 \
4922     pains Nasal_Congestion Runny_Nose Diarrhea      NaN      NaN

      Symptom_11 Symptom_12 Symptom_13 Symptom_14 Symptom_15 Symptom_16 \
4922         NaN         NaN         NaN         NaN         NaN         NaN

      Symptom_17
4922         NaN
covid
```

Predicted Disease

In [4]:

```
print(out['Disease'].unique()[0])
```

```
covid
```