



專題 白酒品質分類預測器

White Wine Classifier

Wei-Ning Yang

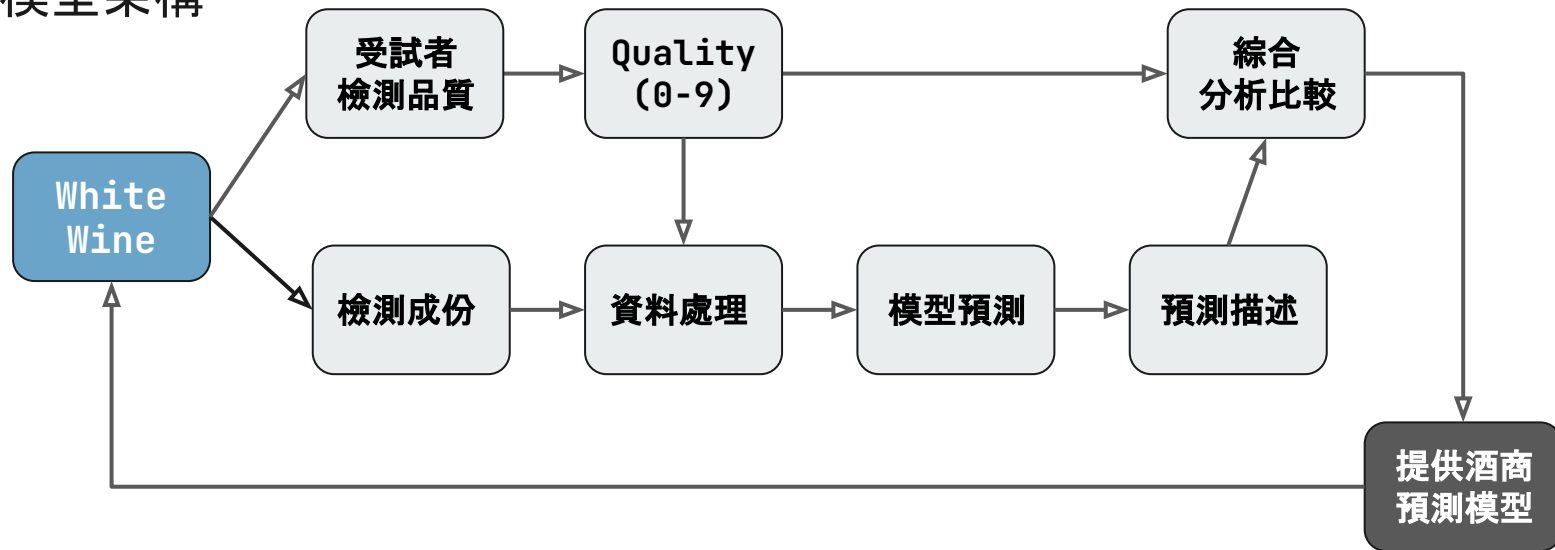


主旨

酒類的香氣主要是來自於醇類、醛類、酸類、酯類以及雜醇油交互影響的結果。 所以這些化合物之含量及濃度多寡對於酒的嗅覺及口感官能品評上是具有正面之意義。

此專題之目的為建立一個以酒的質地、風味、氣味為基礎之模型，用以預測酒類品質，為釀酒廠商提供一套可預測客戶偏好之工具，助酒商研發更符合市場偏好的商品。

模型架構



	A	B	C	D	E	F	G	H	I	J	K	L
1	fixed acid	volatile acid	citric acid	residual sugar	chlorides	free sulfur	total sulfur	density	pH	sulphates	alcohol	quality
2	7	0.27	0.36	20.7	0.045	45	170	1.001	3	0.45	8.8	6
3	6.3	0.3	0.34	1.6	0.049	14	132	0.994	3.3	0.49	9.5	6
4	8.1	0.28	0.4	6.9	0.05	30	97	0.9951	3.26	0.44	10.1	6
5	7.2	0.23	0.32	8.5	0.058	47	186	0.9956	3.19	0.4	9.9	6
6	7.2	0.23	0.32	8.5	0.058	47	186	0.9956	3.19	0.4	9.9	6
7	8.1	0.28	0.4	6.9	0.05	30	97	0.9951	3.26	0.44	10.1	6
8	6.2	0.32	0.16	7	0.045	30	136	0.9949	3.18	0.47	9.6	6
9	7	0.27	0.36	20.7	0.045	45	170	1.001	3	0.45	8.8	6
10	6.3	0.3	0.34	1.6	0.049	14	132	0.994	3.3	0.49	9.5	6
11	8.1	0.22	0.43	1.5	0.044	28	129	0.9938	3.22	0.45	11	6
12	8.1	0.27	0.41	1.45	0.033	11	63	0.9908	2.99	0.56	12	5
13	8.6	0.23	0.4	4.2	0.035	17	109	0.9947	3.14	0.53	9.7	5
14	7.9	0.18	0.37	1.2	0.04	16	75	0.992	3.18	0.63	10.8	5
15	6.6	0.16	0.4	1.5	0.044	48	143	0.9912	3.54	0.52	12.4	7
16	8.3	0.42	0.62	19.25	0.04	41	172	1.0002	2.98	0.67	9.7	5
17	6.6	0.17	0.38	1.5	0.032	28	112	0.9914	3.25	0.55	11.4	7
18	6.3	0.48	0.04	1.1	0.046	30	99	0.9928	3.24	0.36	9.6	6
19	6.2	0.66	0.48	1.2	0.029	29	75	0.9892	3.33	0.39	12.8	8
20	7.4	0.34	0.42	1.1	0.033	17	171	0.9917	3.12	0.53	11.3	6
21	6.5	0.31	0.14	7.5	0.044	34	133	0.9955	3.22	0.5	9.5	5
22	6.2	0.66	0.48	1.2	0.029	29	75	0.9892	3.33	0.39	12.8	8
23	6.4	0.31	0.38	2.9	0.038	19	102	0.9912	3.17	0.35	11	7
24	6.8	0.26	0.42	1.7	0.049	41	12					
25	7.6	0.67	0.14	1.5	0.074	25	16					
26	6.6	0.27	0.41	1.3	0.052	16	14					
27	7	0.25	0.32	9	0.046	56	24					
28	6.9	0.24	0.35	1	0.052	35	14					
29	7	0.28	0.39	8.7	0.051	32	14					
30	7.4	0.27	0.48	1.1	0.047	17	132	0.9914	3.19	0.49	11.6	6
31	7.2	0.32	0.36	2	0.033	37	114	0.9906	3.1	0.71	12.3	7
32	8.5	0.24	0.39	10.4	0.044	20	142	0.9974	3.2	0.53	10	6
33	8.3	0.14	0.34	1.1	0.042	7	47	0.9934	3.47	0.4	10.2	6

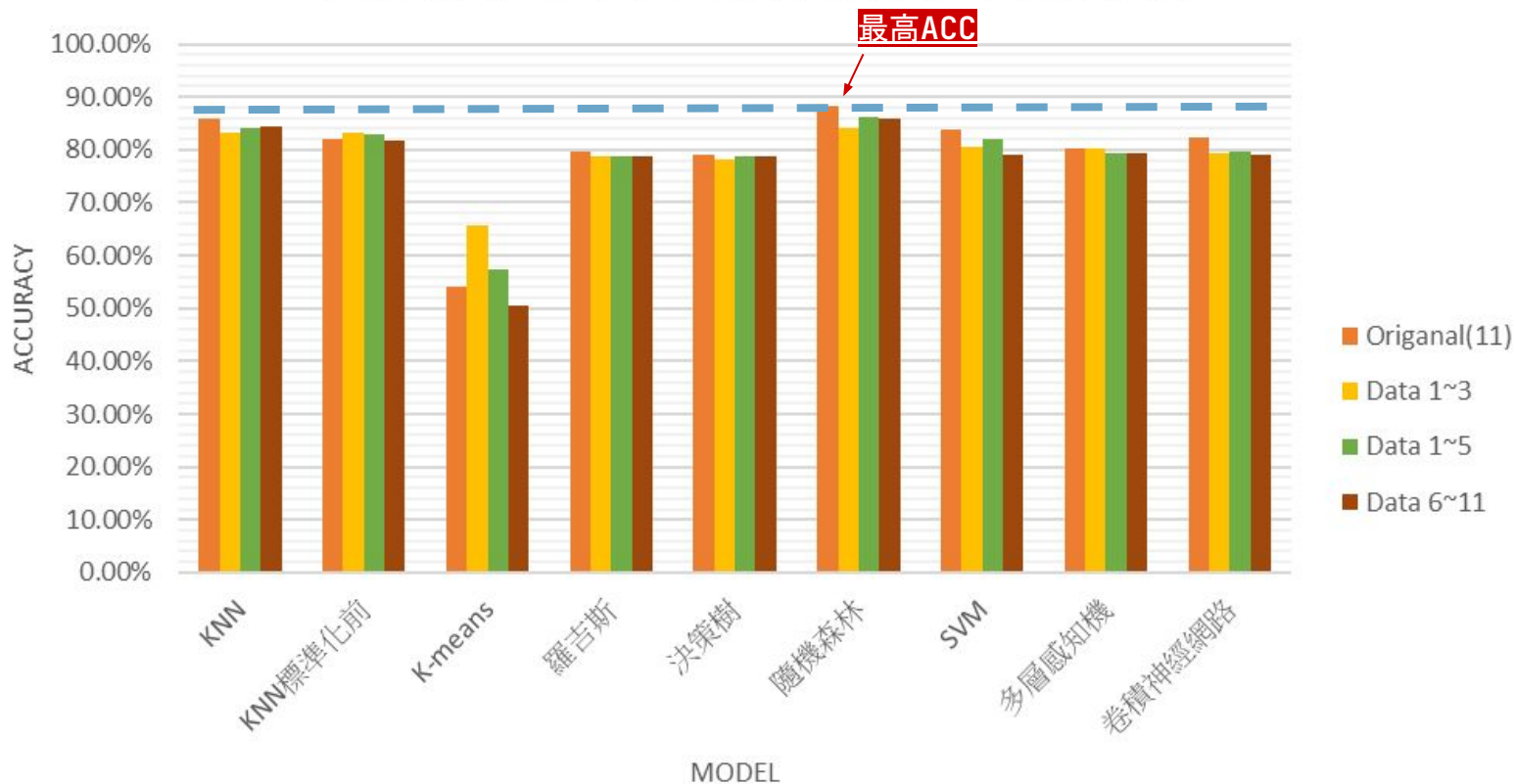
- Data來源 : [Kaggle-White Wine Quality](#)
- 總數4898筆資料
- quality 6.5以上為好酒，以下不是好酒

模型選擇依據

Accuracy	KNN	K-means	羅吉斯	決策樹	隨機森林	SVM	多層感知機	卷積神經網路
Original (11)	85.78%	54.12%	79.73%	79.05 %	<u>88.23%</u>	83.80%	80.40%	82.23%
Data 1~3 (3)	83.27%	65.78%	78.91%	78.23 %	84.08%	80.54%	80.13%	79.36%
Data 1~5 (5)	84.15%	57.21%	78.71%	78.64 %	86.26%	82.04%	79.45%	79.59%
Data 6~11 (6)	84.42%	50.65%	78.64%	78.71 %	85.85%	79.21%	79.31%	79.21%

- 依據重要性分析，採不同數量之Function數量去做預測做比較
- 結果顯示，整體而言全部function做預測準確率最高，其中又以隨機森林模型準確率最高。

各個模型對不同資料集準確度分析



2. 小工具 LOGO



+ 須將PNG檔
轉為ICO檔

3. 顯示的內容

```
label_greeting = Label(text="Please enter values below!", font=FONT)
label_greeting.grid(row=0, column=0, columnspan=2, sticky="EW", pady=20)
label1 = Label(text="fixed acidity")
label1.grid(row=1, column=0)
label2 = Label(text="volatile acidity")
label2.grid(row=2, column=0)
label3 = Label(text="citric acid")
label3.grid(row=3, column=0)
label4 = Label(text="residual sugar")
label4.grid(row=4, column=0)
label5 = Label(text="chlorides")
label5.grid(row=5, column=0)
label6 = Label(text="free sulfur dioxide")
label6.grid(row=6, column=0)
label7 = Label(text="total sulfur dioxide")
label7.grid(row=7, column=0)
label8 = Label(text="density")
label8.grid(row=8, column=0)
label9 = Label(text="pH")
label9.grid(row=9, column=0)
label10 = Label(text="sulphates")
label10.grid(row=10, column=0)
label11 = Label(text="alcohol")
label11.grid(row=11, column=0)
```

1. 小工具標題

Wine Quality Classifier

```
window.title("Wine Quality Classifier")
```

4. 使用者輸入 feature值

限制使用者輸入內容
僅可輸入正負浮點數

```
#限制輸入正負浮點數
def validate(s: str) -> bool:
    return re.match(r"[-+]?[0-9]*\.?[0-9]*", s).group() == s
vcmd = (window.register(validate), '%P')

inputs = []
for j in range(1,12):
    input = Entry(validate='key', validatecommand=vcmd)
    input.grid(row=j, column=1, sticky="EW")
    inputs.append(input)
    j += 1
```

```
def classify():
    vals = []
    for i in inputs:
        if i.get() == '':
            x=0
            vals.append(x)
        else:
            x = float(i.get())
            vals.append(x)
```

若使用者未輸入
則判定為0
仍可進行預測

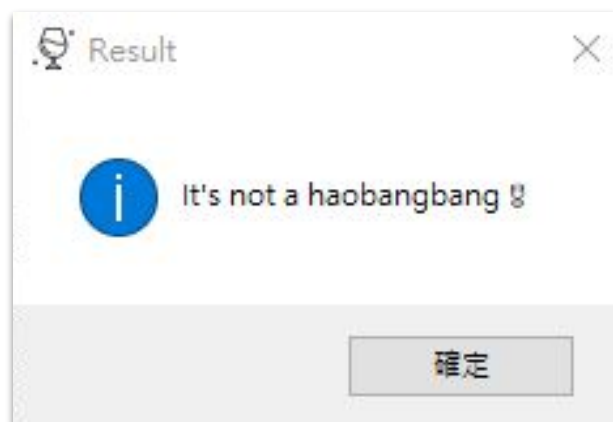
Please enter values below!

fixed acidity
volatile acidity
citric acid
residual sugar
chlorides
free sulfur dioxide
total sulfur dioxide
density
pH
sulphates
alcohol

Classify

5. 執行預測案紐

```
button = Button(text="Classify", command=classify)
button.grid(row=12, columnspan=2, pady=20)
```



6. 預測結果顯示

```
with open("classifier.pkl", "rb",) as f:
    knn = pickle.load(f)
    pred = knn.predict(x)
    if pred[0] == 1:
        messagebox.showinfo(title="Result", message="It's a haobangbang 🙄")
    else:
        messagebox.showinfo(title="Result", message="It's not a haobangbang 🙄")
vals.clear()
for i in inputs:
    i.delete(0, END)
```




THE END

Thank you for your patience