**Naivný Bayes**

**Úvod**

V mojom projekte je účelom klasifikácia jedlých a nejedlých húb na základe ich atributov. Obsah:

1) Implementácia Bayes klasifikátora

2) Návod na použitie aplikácie založenej na tomto klasifikátore

**Implementácia klasifikátora**

Podľa mojej implementácie je model Bayes klasifikátora rozdelený do 3 hlavných funkcií:

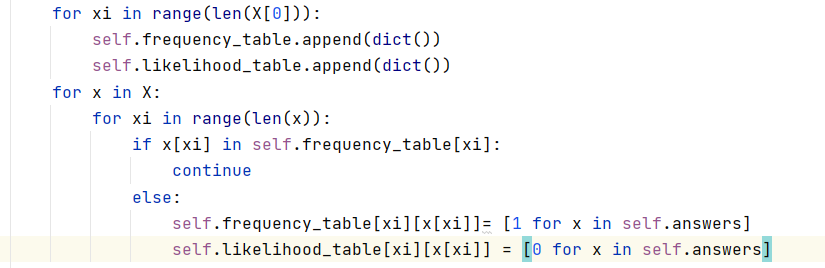
* Fit
* Predict
* Test

A a tiež 2 dodatočné:

* Save
* Load

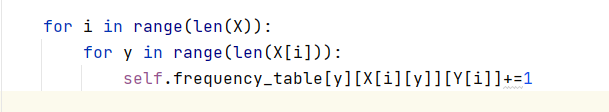
**1)Fit**

Hlavným účelom algoritmu klasifikátora je vytvorenie pravdepodobnostnej tabuľky na základe frekvenčnej tabuľky.



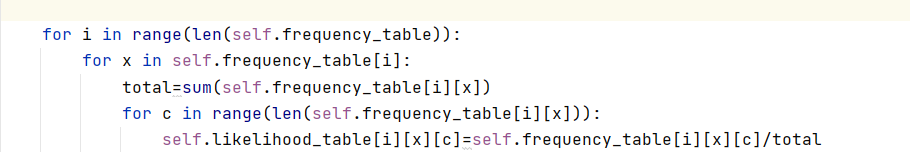
Obrazovka 1: Inicializácia tabuľek

Pred vyplnením tabuľky pravdepodobnosti algoritmus vyplní tabuľku frekvencií, obr.2



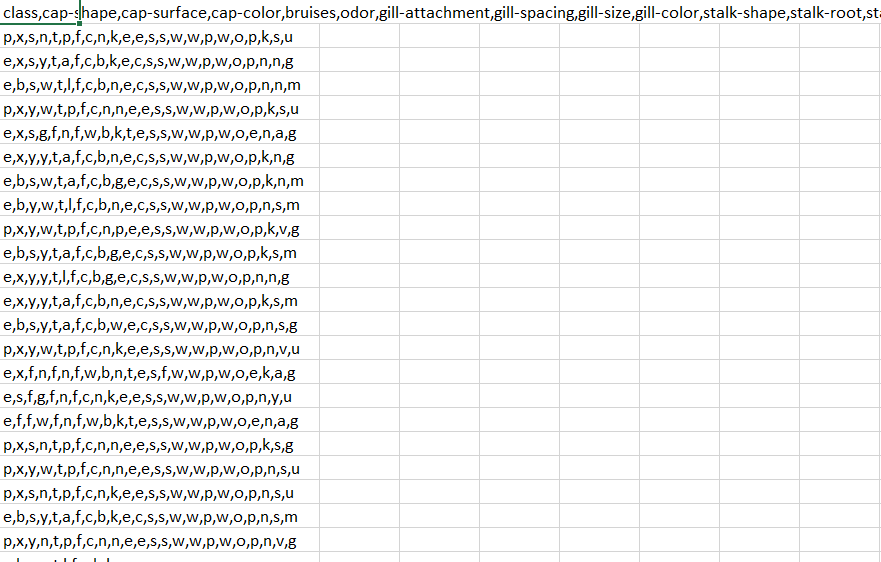
Obrazovka 2: Vyplnenie tabuľky frekvencií

Potom zostavíme samotnú tabuľku pravdepodobností, obr. 3



Obrazovka 3: Vyplnenie tabuľky pravdepodobnosti

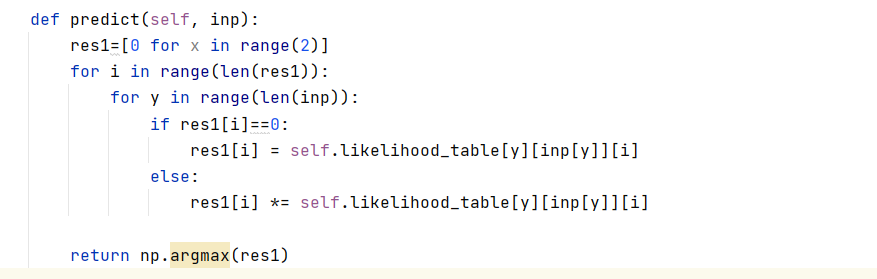
Klasifikátor bol trénovaný na datasete mushrooms.csv, <https://www.kaggle.com/uciml/mushroom-classification>, obr. 4



Obrazovka 4: mushrooms.csv

**2)Predict**

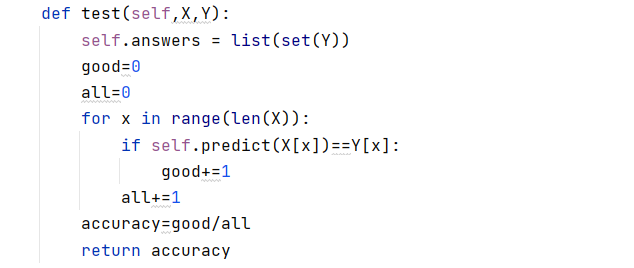
Funkcia Predict () používa Bayesovu teorému na výpočet pravdepodobnosti [príslušnosťi vstupu do každej tried,](https://webslovnik.zoznam.sk/slovensko-rusky/?s=pr%C3%ADslu%C5%A1nos%C5%A5)a potom pomocou funkcie argmax () vyberie triedu s najväčšou pravdepodobnosťou.



Obrazovka 5: predict()

**3)Test**

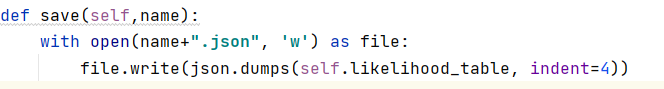
Funkcia test() nav stupe berie vstupné údaje, na základe ktorých bude model testovaný, na klasifikáciu týchto údajov používa funkciu predict() a potom vypočítava presnosť modelu porovnaním správnych odpovedí s získanými.



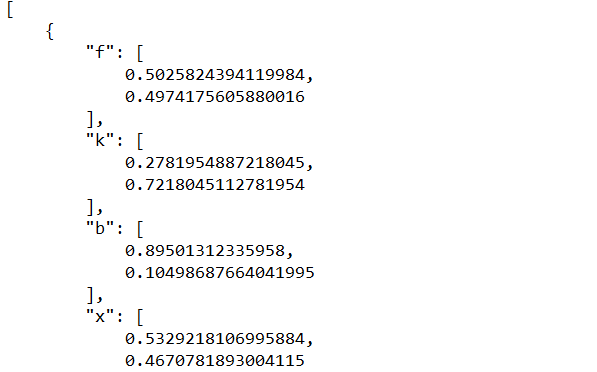
Obrazovka 6: test()

**4)Save**

Funkcia save () uloží tabuľku pravdepodobností získanú po učeniu ako súbor .json.



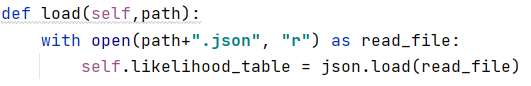
Obrazovka 7: save()



Obrazovka 8: Časť uloženej tabuľky pravdepodobnosti vo formáte .json

**4)Load**

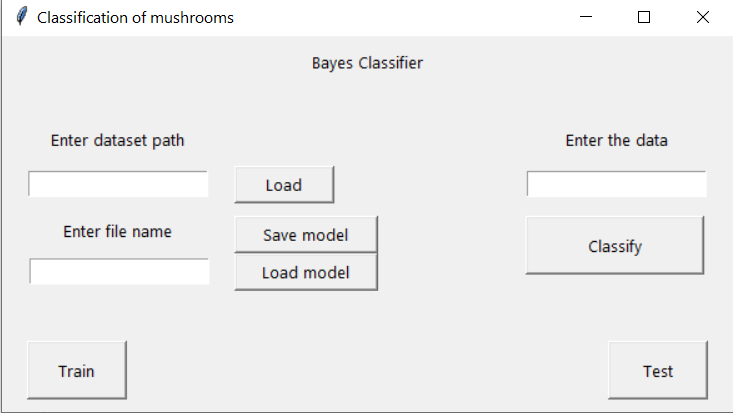
Funkcia load () načíta uloženú pravdepodobnostnú tabuľku do Bayesovho modelu, vďaka čomu nie je potreba opakovaného učenia



Obrazovka 9: load()

**Návod na použitie aplikácie**

Aplikácia bola napísaná v Pythone pomocou knižnice **TKinter**, pomocou tejto aplikácie môžete trénovať Bayesov klassificator a použiť ho na klasifikáciu údajov používateľov, konkrétne na určenie typu húb (jedovatých alebo jedlých).



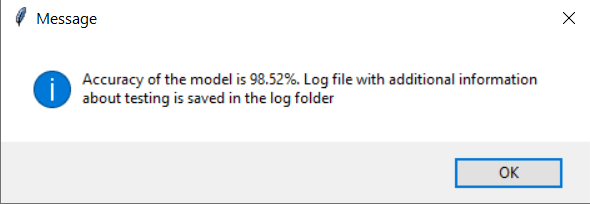
Obrazovka 10: interface aplikácií

1. **Učenie a testovanie modelu**

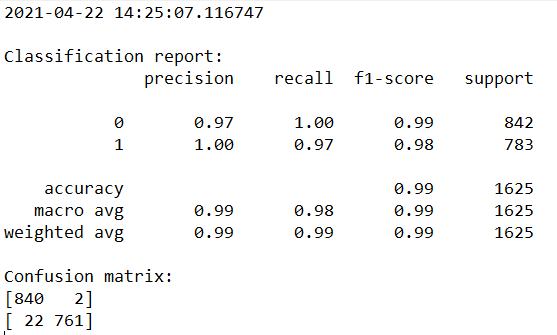
Návod na učenie modelu:

1. Do príslušného poľa zadajte cestu k datasetu a kliknite na tlačidlo **Load**
2. Stlačte tlačidlo **Train**

Pre otestovanie modelu po učeniu je potrebné kliknúť na tlačidlo **Test**, vďaka tomu sa presnosť modelu zobrazí vo vyskakovacom okne, obr. 11 a v priečinoku **Log** sa objaví log-súbor, v ktorom budú podrobnejšie informácie o testovaní Obr. 12. Ak chcete, po učeniu modelu, ho môžete uložiť zadaním názvu modelu do príslušného poľa a kliknutím na tlačidlo **Save** a potom ho načítať rovnakým spôsobom iba kliknutím na tlačidlo **Load**.



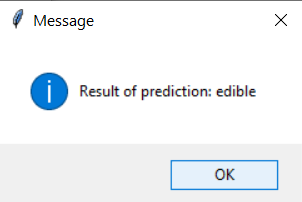
Obrazovka 11



Obrazovka 12: Obsah log-suboru

1. **Používanie modelu**

Aby ste mohli klasifikovať vstupy používateľov, musíte najskôr trénovať model alebo načítať hotový, potom zadajte údaje do príslušného poľa, príklad vstupu: „b,s,w,t,l,f, c,b,n,e,c,s,s,w,w,p,w,o,p,n,n,m “. Potom kliknite na tlačidlo **Classify**, výsledok predikcie sa zobrazí v vyskakovacom okne, obr. 13.



Obrazovka 13: Výsledok predikcie