گزارش تمرین چهارم یادگیری ماشین

نام : حامد خشه چی

شماره دانشجویی : ۹۷۲۳۴۹۴

جدول زیر را پر کنید.

Best word of	Class1	Class2	Class3	Class4	Class5	Class6
1	Graphics	Forsale	Autos	Encryption	Rutgers	**Guns**
2	Ompass	Misc	Rec	Key	Athos	**Politics**
3	op	Computer	Car	Crypt	Christian	**Gun**
4	Polygon	Sale	Cars	Sci	Od	**Ifas**
5	Raphics	Shipping	Dealer	Security	Esus	**Gnv**
6	Animation	Music	Engine	Org	Religion	**Firearms**
7	Mage	Sunysb	Vw	Privacy	Soc	**Ufl**
8	Cols	Clone	Cactus	Eff	Hedrick	**Talk**
9	Rows	Condition	Cdac	Escrow	Aramis	**Weapons**
10	Au	Oit	Ole	Technology	Igor	**Firearm**
11	Acsc	Unc	Aturn	Lipper	Pproved	**Utk**
12	Int	Pc	W	Cryptography	Hristians	**Weapon**
13	Row	Marketplace	Tires	Government	Hrist	**Ulowell**
14	Image	Gibbs	Ntegra	Keys	Hristianity	**Iftccu**
15	Points	Sking	Fuel	Answers	Truth	**Utkvm**
16	Eality	Ba	Kocrsv	Enforcement	Faith	**Laws**
17	Raf	Ale	Brandeis	Hip	Hristian	**Police**
18	Irtual	Pchang	Pockets	Dministation	Ible	**Crime**
19	Color	96	Boyle	Secure	Dziuxsolim	**Defense**
20	ys	sbcs	radar	data	church	**Arms**

graphics', 'ompass', 'op', 'polygon', 'raphics', 'animation', 'mage', 'cols', 'rows', 'au', 'acsc', 'int', 'row', ']
['image', 'points', 'eality', 'raf', 'irtual', 'color', 'ys

forsale', 'misc', 'computers', 'sale', 'shipping', 'music', 'sunysb', 'clone', 'condition', 'oit', 'unc', 'pc', ']
['marketplace', 'gibbs', 'sking', 'ba', 'ale', 'pchang', '96', 'sbcs

autos', 'rec', 'car', 'cars', 'dealer', 'engine', 'vw', 'cactus', 'cdac', 'ole', 'aturn', 'W', 'tires', 'ntegra', 'fuel', ']
['kocrsv', 'brandeis', 'pockets', 'boyle', 'radar

encryption', 'key', 'crypt', 'sci', 'security', 'org', 'privacy', 'eff', 'escrow', 'technology', 'lipper', ']
''cryptography', 'government', 'keys', 'answers', 'enforcement', 'hip', 'dministration', 'secure', 'data'[

rutgers', 'athos', 'christian', 'od', 'esus', 'religion', 'soc', 'hedrick', 'aramis', 'igor', 'pproved', 'hristians', ']
''hrist', 'hristianity', 'truth', 'faith', 'hristian', 'ible', 'dziuxsolim', 'church'[

guns', 'politics', 'gun', 'ifas', 'gnv', 'firearms', 'ufl', 'talk', 'weapons', 'firearm', 'utk', 'weapon', 'ulowell', ']
''iftccu', 'utkvm', 'laws', 'police', 'crime', 'defense', 'arms'[

خروجی های ما بوده اند برای هر کلاس

که توسط نوشتن رگولار اکسپرشن روبرو به دست آمده اند

```
(?<!\d)[<>(),|*-_.?$!:\'`@"]|[<>()$,.!-_?|*`\':@"](?!\d)|\s+
```

که یعنی هر چی به جز وایت اسپیس ها و علامات خاص بود جدا کن و به ما بده که کلمات اند که فایل وکا نیز بر اساس همین ها به دست آمده و دقت نیو ما و وکا هردو یک بوده است

۱,۰

۱,۰

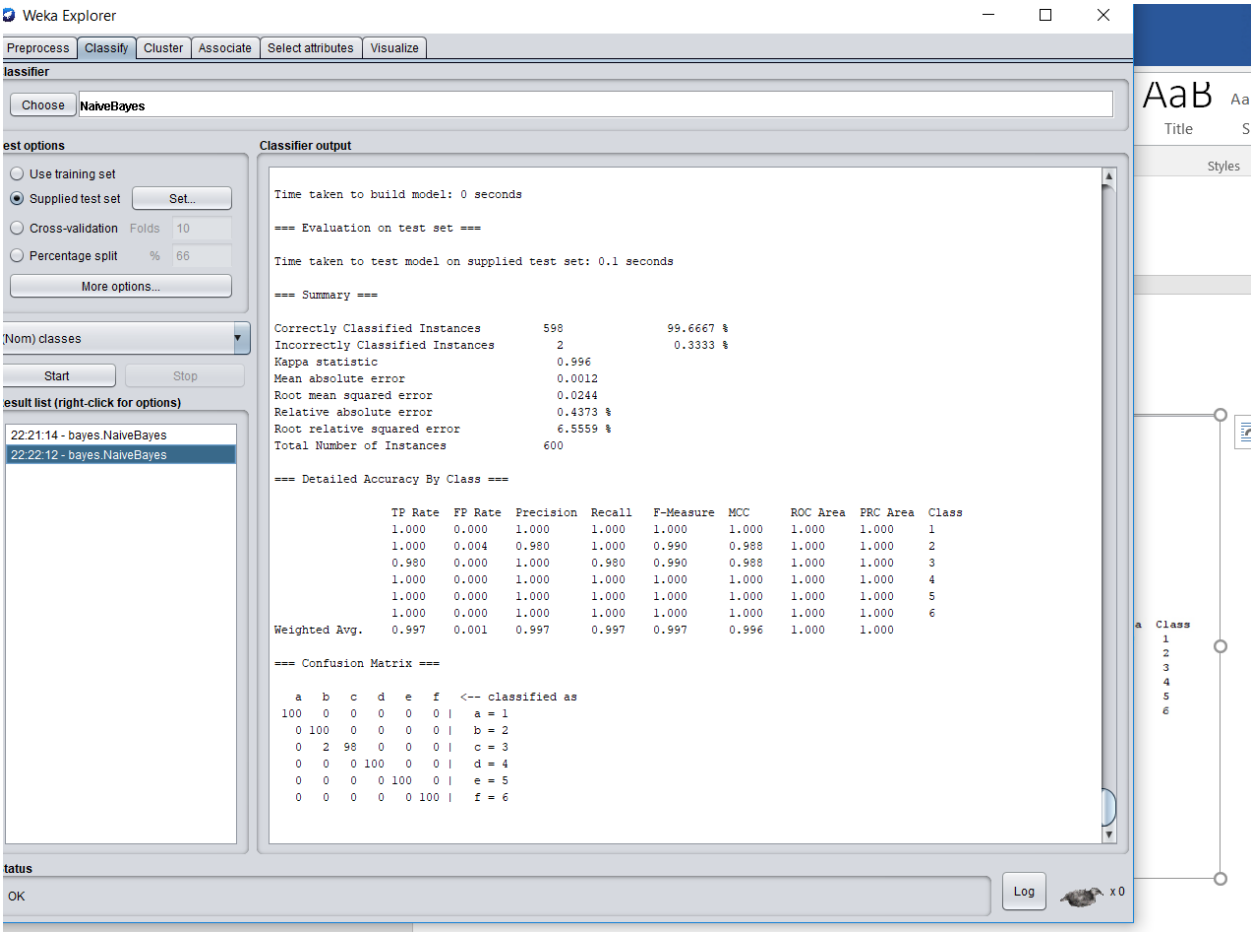
۱,۰

۱,۰

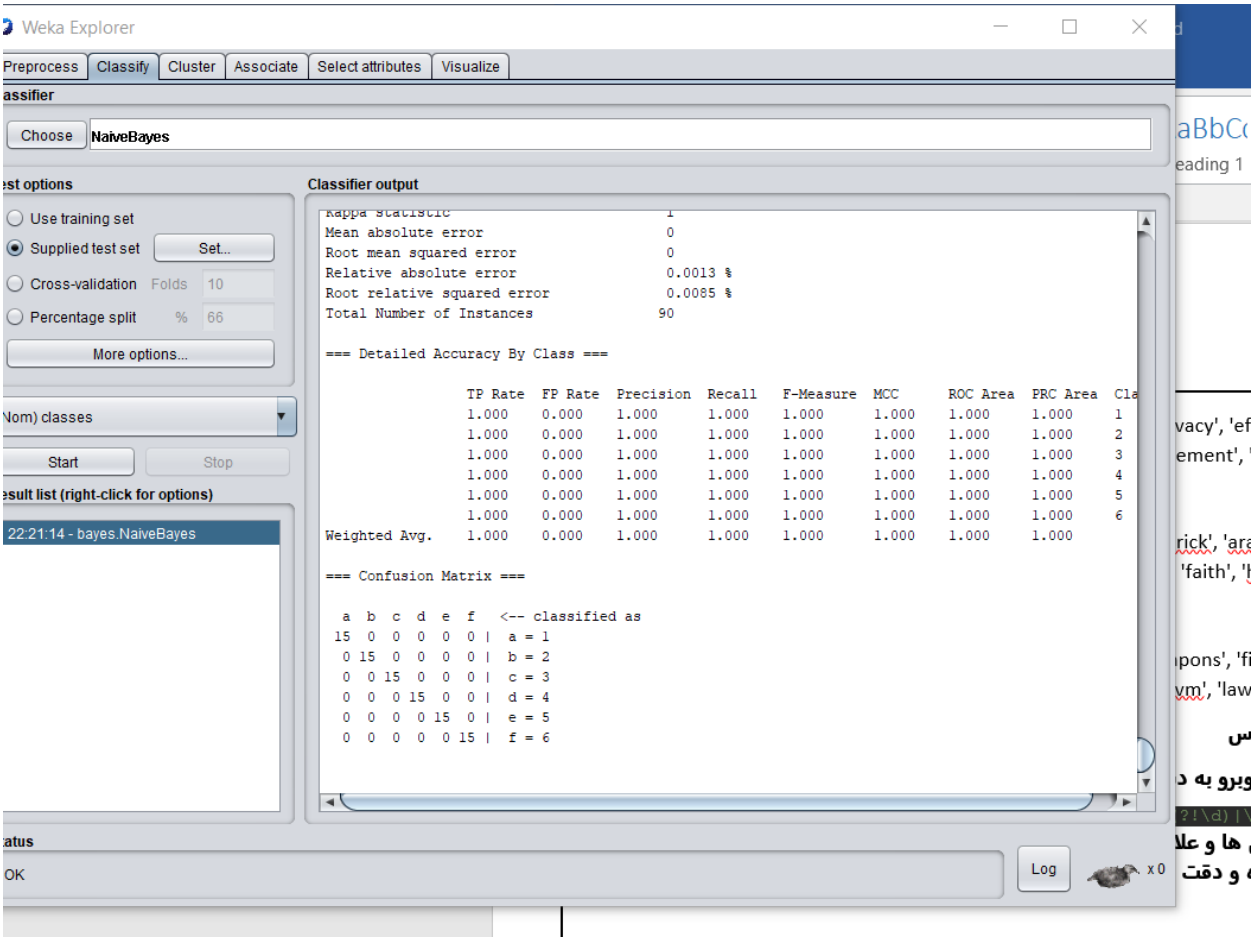
۱,۰

۱,۰

خروجی classifier naïve bayes **برای داده های آموزشی (خروجی نرم افزار** weka)



خروجی classifier naïve bayes **برای داده های تست (خروجی نرم افزار** weka)



و برای نوشتن کد از فایل کمکی استفاده کردم که یعنی من هم مقدار موثر را طبق همان فرمول حساب می کنم. که بیشترین تکرار در خود که در باقی نیست باشد!