

```
1 /Users/MikiWong/anaconda3/envs/scikitlearn/bin/python3 /  
2   Users/MikiWong/PycharmProjects/amazonscraper/classification  
3     /classification2.py  
4  
5  
6  
7  
8  
9  
10 Shapes (X_Train, Y_train): (10093, 50969) (10093,)  
11 Shapes (X_Test, Y_test): (2524, 50969) (2524,)  
12  
13 =====TRAINING NAIVES BAYES:MULTINOMIAL=====  
14             precision      recall    f1-score   support  
15  
16 business      0.86675  0.99022  0.92438    716  
17 cook          0.99411  0.93358  0.96289    542  
18 literature    0.98188  0.91834  0.94904    649  
19 sports         0.96949  0.92707  0.94780    617  
20  
21 avg / total   0.94882  0.94414  0.94472    2524  
22  
23 InterAnnotator Agreement score: 0.925093405924  
24 =====TRAINING SVM.LINEAR SVC=====  
25 Time taken to train: 0.4302341938018799  
26             precision      recall    f1-score   support  
27  
28 business      0.96419  0.97765  0.97087    716  
29 cook          0.98881  0.97786  0.98330    542  
30 literature    0.96224  0.98151  0.97178    649  
31 sports         0.99333  0.96596  0.97946    617  
32  
33 avg / total   0.97610  0.97583  0.97587    2524  
34  
35 InterAnnotator Agreement score: 0.96765740766  
36  
37 ===SAMPLE PREDICTIONS FROM SVM====  
38  
39 Sample description: if seem like nashville everywhere day  
gq recently declare nowville become music hotspot country  
rock but hot music scene food scene even hotter nbsp in  
nashville eats 100 mouthwatering recipe reveal food lover  
head south nashville hot chicken buttermilk biscuit pull  
pork sandwich cornmeal crust catfish chowchow fry green  
tomato chess pie author jennifer justus whip classic  
pimento cheese fry chicken also include dish twist  
traditional southern fare curry black chickpea catfish taco  
and alongside recipe jennifer share story nashville people
```

39 music history food make special
40 Label: cook
41 Predicted class: ['cook']
42
43 Sample description: love trader joe this new independent cookbook feature recipes use ingredient trader joe by combine trader joe unique product fresh ingredient deana wona create clever shortcuts quick easy gourmet meal delicious excite the recipe book treat trader joe like prep kitchen use great selection unique sauce mixtures prepped item make flavorful natural homemade food snap many recipe vegetarian easily make vegetarian ethnic dish like saag paneer lasagna scatter throughout well classic comfort food like comfy chicken pot pie crowd please recipe include peanutty sesame noodle black bean soup macho nacho seafood paella curry chicken pita wilt spinach attitude honey i eat chocolate bread pudding all mix up margaritas people know cook want cook appreciate bachelor quicky section feature frozen ready heat selection match create complete impressive menu with full color photograph every recipe wine suggestion humorous personal story cook tip sprinkle throughout collection must trader joe fan
44 Label: cook
45 Predicted class: ['cook']
46
47 Sample description: you likely hear coffee professional local coffee shop roaster talk different excite flavor note bean produce find disappointed brew bean home taste coffee just like taste wine develop palate pick subtle note present high quality coffee bean take time practice yield delicious dividend this book give crash course taste coffee it include information note often expect different grow region roast level advice train palate information taste wheel use professional tip professional taster brew coffee best taste layer flavor some coffee quite complex aromas flavor subtle fleet this book help learn taste like barista
48 Label: cook
49 Predicted class: ['cook']
50
51 Sample description: an new collection delicious simple vegan dish use easy find readily available ingredient go vegan never easier the author the mediterranean vegan kitchen show reader use readily available ingredient create healthy delicious simple low cost elegant family vegan meal include more 225 meat free egg free dairy free recipes appetizer soup salad main course side dish beverage desserts no expensive suspect meat egg dairy substitute helpful preparation tip comprehensive nutritional analysis every recipe
52 Label: cook

- 53 Predicted class: ['cook']
54
55 Sample description: original hardcover seal cookbook year
the james beard award copyright 1991 504 page size 7 x 9 1
2 technique sauce make change radically year both
professional chef serious amateur cook eager invent new
taste combination improve upon traditional method set make
sauce healthy less rich well until book however rely
trial error produce innovative new sauce single source
systematically describe ingredient method principle use
make classic sauce modern variation this comprehensive
volume provide complete analysis classic contemporary
sauce construct flavored sauce explain ingredient sauce
combine over 300 recipe comprise classic modern sauce
delicious first course entree preparation incorporate
sauce make technique describe you also find practical
discussion equipment ingredient commonly use sauciers
advice use different kind cookware tool machinery serve
utensil select high quality ingredient key prepare
excellent sauce learn choose use dairy product herbs spice
win spirit mushroom truffle chocolate pork product
ingredient best advantage a comprehensive glossary term
flow chart illustrate major classical sauce family method
preparation round reference one excite challenge area cook
56 Label: cook
57 Predicted class: ['cook']
58
59 Sample description: the owner char broil smoker recipe
cookbook for smoke pork beef poultry fish amp wild game
smoker recipe cookbook ever need beautiful full color
picture the owner char broil smoker recipes cookbook
ultimate guide char broil smoker recipes whether barbecue
grill the smoker recipes satisfy beginner advance chef
this smoker cookbook make smoke favorite meat simple easy
prepare the owner char broil smoker recipe cookbook for
smoke pork beef poultry fish amp wild game fill
mouthwatering smoker recipe learn secret tip trick smoke
meat like professional fast meat temperature chart savior
aroma flavor never think possible rub marinade
professional chef disclose learn variety wood type flavor
informative wood chart be sure purchase owner char broil
smoker recipe for smoked vegetable nutritional delicious
side dish enjoy smoked meat
60 Label: cook
61 Predicted class: ['cook']
62
63 Sample description: the petrakis family live small greek
seaside village plaka just coast tiny island spinalonga
nation leper colony locate place haunt four generation
petrakis woman there eleni ripped husband two young

63 daughter send spinalonga 1939 daughter maria find joy everyday dutifully care father anna wild child hungry passion life anywhere plaka and finally alexis eleni great granddaughter visit modern day greece unlock family past a richly enchant novel life love unfold backdrop mediterranean world war ii the island enthrall story dream desire secret desperately hide leprosy touch unforgettable family

64 Label: literature

65 Predicted class: ['literature']

66

67 Sample description: in 1 new york time bestselling author j r ward thrill finale bourbon king series bradford family dynasty teeter edge collapse murder patriarch shock arrest nbsp at first death william baldwine head bradford family rule suicide but eldest son swear enemy edward come forward confess fact murder now police custody edward mourn disintegration family loss freedom woman leave behind his love sutton smythe person ever truly cared ceo bradford bourbon company big competitor relationship impossible and reality jail time edward face nbsp lane baldwine suppose remain role playboy forever big brother edward shadow instead become new head family company convince edward cover someone else lane true love lizzie king go trail killer discover secret devastating game change nbsp a lane rush discover truth sutton find irresistibly draw edward spite circumstance live everyone easterly never for good others could tragedy beyond imagine only one thing certain love survive thing even murder praise j r ward bourbon king series nbsp a sinful indulgence the gorgeous kentucky backdrop sweep away susan wiggs nbsp breathless fun lisa gardner

68 Label: literature

69 Predicted class: ['literature']

70

71 Sample description: a deluxe hardcover edition pioneer cyberpunk novel predict obsession internet part penguin galaxy collectible series six sci fi fantasy classic feature series introduction neil gaimanwinner aiga design observer 50 book 50 cover competition before internet commonplace william gibson show u matrix world within world representation every byte data cyberspace henry dorsett case sharp data thief matrix ex employer cripple nervous system now new employer recruit last chance run unthinkably powerful artificial intelligence with mirror eye girl street samurai rid shotgun ready silicon quick bleakly prophetic adventure up ante entire genre fiction penguin galaxy nbsp six great masterworks science fiction fantasy dazzle collector worthy hardcover edition feature series introduction 1 new york time bestselling author

71 neil gaiman penguin galaxy represent constellation achievement visionary fiction light way toward knowledge universe from historical legend mythic future monument world building mind bend dystopias touchstone human invention storytelling ingenuity transport million reader distant realm continue generation chart frontier imagination nbsp the once future king t h white stranger strange land robert a heinlein dune frank herbert 2001 a space odyssey arthur c clarke the left hand darkness ursula k le guin neuromancer william gibson for seventy year penguin lead publisher classic literature english speaking world with 1 700 title penguin classic represent global bookshelf best work throughout history across genre discipline reader trust series provide authoritative text enhance introduction note distinguish scholar contemporary author well date translation award win translator

72 Label: literature

73 Predicted class: ['literature']

74

75 Sample description: from new york time bestselling author kristin hannah come powerful novel love loss magic friendship in turbulent summer 1974 kate mularkey accept place bottom eighth grade social food chain then amazement cool girl world move across street want friend tully hart seem beauty brain ambition on surface opposite two people kate doom forever uncool love family mortifies every turn tully steep glamour mystery secret destroying they make pact best friend forever summer end become tullyandkate inseparable so begin kristin hannah magnificent new novel span three decade play across ever change face pacific northwest firefly lane poignant powerful story two woman friendship becomes bulkhead life from begin tully desperate prove worth world abandon mother early age longs love unconditionally in glitter big hair era eighties look men fill void soul but button ninety television news captivate she follow blind ambition new york around globe find fame success loneliness kate know early life nothing special throughout college pretend driven need success really want fall love child live ordinary life in quiet way kate drive tully what know wife mother change lose sight want and much envy famous best friend for thirty year tully kate buoy life weather storm friendship jealousy anger hurt resentment they think survive single act betrayal tear apart put courage friendship ultimate test firefly lane anyone ever drink boone farm apple wine listen abba fleetwood mac more coming age novel story generation woman bless cursed choice it promise secret betrayal and ultimately one person really truly know know power hurt heal firefly lane story never forget one want

75 pas best friend
76 Label: literature
77 Predicted class: ['literature']
78
79 Sample description: the glass menagerie read text author prefer available new direction paperback edition a new introduction editor the tennessee williams annual review robert bray reappraises play half century new york drama critic award this edition the glass menagerie also include williams essay impact sudden fame struggle writer the catastrophe success well short section williams production note
80 Label: literature
81 Predicted class: ['literature']
82
83 Sample description: winner 2000 pulitzer prize drama in ponder issue intimacy passion pleasure dolor conjugal routine margulies write relationship intelligence spiky humor comedy drama...becomes something quite wonderful – william tynan time a true feast dinner friends sober wise extremely funny...mr margulies write elegance wit intimacy trust privacy toll time monogamy sexual passion matter easily express dramatic dialogue –vincent canby new york time the complexity marriage friendship dissect wit intelligence dinner friend at tender disturbing... drawn intriguing psychological complexity feel quietly real margulies depict vagary spousal relationship humor pathos insightful manner undoubtedly spur meaningful conversation couple –frank scheck hollywood reporterfilled humor warmth insight wisdom dinner friend modern day masterpiece destruction today marriage through margulies flawless use language ability convey true dialogue characterization watch two couple closest friend go wrench breakup not create vivid detail marriage decline also brilliantly depict couple close friends new mirror marriage send whirlwind raw emotion self reflection donald margulies receive 2000 pulitzer prize drama dinner friends the play receive numerous award include american theatre critic association new play award dramatists guild hull warriner award lucille lortel award outer critic circle award drama desk nomination produce united state around world in addition adaptation god vengeance many play include collect story the country house sight unseen the model apartment the loman family picnic what wrong this picture time stand still mr margulies currently live wife son new haven connecticut teach playwriting yale university
84 Label: literature
85 Predicted class: ['literature']
86
87 Sample description: clair bee 1896 1983 hugely successful

- 87 basketball coach rider college long island university 412
87 record career derail 1951 point shave scandal in trial
sent star player sherman white prison judge excoriate bee
create morally lax culture contribute player involvement
gamble to certain extent bee agree judge scold conclude
coach include become driven succeed court lose sight
educational role sport play his coach career effectively
bee launched effort reform ills saw college sport page
chip hilton novel young reader he begin series 1948 post
scandal book use teaching tool the book mirror event
gamble scandal bee attempt reform problem plague college
sport he use fiction posit well sport world hop young
reader would construct inhabit the chip hilton book
extremely popular become classic series two million copy
sell date hoop crazy fascinate story clair bee star
character chip hilton way live real fictional intertwine
- 88 Label: sports
- 89 Predicted class: ['sports']
- 90
- 91 Sample description: born 1970s radical challenge
traditional baseball statistic sabermetrics develop new
way understanding many aspect game it practitioner create
new statistical tool revise old way think establish
measure bat average tactic sacrifice bunt even among great
truly great this introduction basic sabermetric analysis
explain concept include normalization peak versus career
performance linear weight run create well popular
calculation like ops on base plus slug whip walk hit per
inning pitch pf park factor others increasingly use
baseball fan
- 92 Label: sports
- 93 Predicted class: ['sports']
- 94
- 95 Sample description: since nick saban hire head coach
alabama 2007 crimson tide establish record excellence
unparalleled college football modern era feature stunning
action photography five championship campaign expert
analysis author christopher walsh bama dynasty inside
story alabama era dominance team path fifth championship
saban this commemorative book take fan alabama first four
championship season well excite journey 2017 lead
quarterback jalen hurt ferocious defense alabama start
season 11 straight win find sweet redemption no 1 clemson
sugar bowl finally top college football world beat sec foe
georgia national championship game relive unforgettable
season must keepsake include feature hurt minkah
fitzpatrick levi wallace ronnie clark roll tide nbsp
- 96 Label: sports
- 97 Predicted class: ['sports']
- 98

- 99 Sample description: the best selling guide essential game big well comprehensive ever solid block sure tackle proper footwork precise route correct coverage angle synchronize body movement basic technique essential win football football skill amp drill guide success explain teach key skill consistent execution player offensive defensive special team position you benefit technique use top college professional team along complete arsenal 114 drill from quarterback rollouts win touchdown block shed maneuver drive stop sack football skill amp drill translate technical instruction field application depict exactly pas pattern route catch block tackle kick perform gridiron champion make practice field football skill amp drill serve great foundation quest title
- 100 Label: sports
- 101 Predicted class: ['sports']
- 102
- 103 Sample description: youth sport fun game anymore what pleasurable experience often mar poor sportsmanship trash talk win cost attitude worst case violence but world cup soccer champion olympic gold medalist brandi chastain solution in it not about bra chastain draws lesson learn phenomenal career experience parent illuminate beautiful game provide creative answer challenge face young athlete parent chastain emphasizes importance develop leadership skill find become role model give back one team community she offer blueprint kid parent alike play fair win lose grace good time
- 104 Label: sports
- 105 Predicted class: ['sports']
- 106
- 107 Sample description: the hot instructor golf sport illustrate present revolutionary guide lower score powerful new approach strategizing decision making management every part short game stan utley breakthrough short game put technique make one sought golf instructor america now offer breakthrough approach golfer skill level course management approach design help golfer crucial stage round golf the art scoring show reader understand way short game handicap overall skill level dictate strategy break pitch chip bunker play put three proficiency category utley present customize technique save shot simply make good decision he lead u round three amateur two 10 handicapper 20 handicapper show improve strategy execution transform player game with behind thescenes pro teach session crash course three hard greenside shot one hundred black white color photograph the art score like get brand new set state art club customize one game premier teacher
- 108 Label: sports

109 Predicted class: ['sports']
110
111 Sample description: discover 6 week fitness plan become
golf fit increase flexibility balance strength endurance
golf course suitable all age fitness level table of
content week 1 get start week 2 stretching cardio week 3
flexibility cardio week 4 strength flexibility week 5
strength 2 flexibility week 6 design your program bonus
pre game warm up routine click buy now button enjoy the
ultimate guide golf player
112 Label: sports
113 Predicted class: ['sports']
114
115 Sample description: it secret hard time even hard black
community approximately 35 percent african american
measurable asset 2009 24 percent household motor vehicle
dennis kimbrow observe weight continue housing credit
crisis disproportionately impact african american
community take sharp look carefully cultivate group
individual scale height success others emulate base seven
year study 1 000 wealthy african americans the wealth
choice offer trove sound surprising advice climb economic
ladder even odds seem stack reader learn business leader
entrepreneurs celebrity like bob johnson spike lee l a
reid herman cain t d jakes tyrese gibson find path wealth
learn money early sacrifice get top role discipline manage
success through story include men woman every stage life
every industry dennis kimbrow show reader · develop wealth
generate mindset habits · commit lifelong learning · craft
goal match passion · make short term sacrifice long term
gain · take calculated risk opportunity present
116 Label: business
117 Predicted class: ['business']
118
119 Sample description: the self direct ira handbook explain
investor alternative retirement plan asset iras these
asset typically include real estate private stock ira llcs
lp llc private equity hedge fund private placement
crowdfunding precious metal the handbook clearly explain
prohibited transaction rule investment structure common
self direct ira investment asset sdira tax pitfall ubit
amp udfi tax the handbook go beyond basic self direct ira
include extensive faq chapter nearly 100 citation case
statute ruling opinion court irs dol it author mat
sorensen experience attorney advised client iras every
major self direct ira custodian speaks frequently industry
association professional group self direct ira topic mat
regularly advises client prohibited transaction rule ira
llc structure real estate private company investment ubit
udfi tax mat also successfully represent self direct ira

- 119 owner irs appeal office u s tax court to view testimonial book additional self direct ira resource please visit author site www sdirahandbook com
- 120 Label: business
- 121 Predicted class: ['business']
- 122
- 123 Sample description: most company change initiative fail yours if read nothing else change management read 10 article feature lead change john p kotter we comb hundred harvard business review article select important one help spearhead change organization hbr 10 must read change management inspire lead change eight critical stagesestablish sense urgencyovercome addiction status quomobilize commitmentsilence naysayersminimize pain changeconcentrate resourcesmotivate change business goodthis collection best selling article include featured article lead change why transformation effort fail john p kotter change through persuasion lead change when business be good an interview samuel j palmisano radical change quiet way tip point leadership a survival guide leader the real reason people win change crack code change the hard side change management why change program don produce change
- 124 Label: business
- 125 Predicted class: ['business']
- 126
- 127 Sample description: take care customer someone else legendary servicegreat customer service concept organization love know yet people consider service receive average best successful company make connection legendary customer service thrive business recognize way employee treat customer directly related way manager treat employee kelsey young optimistic disillusion sale associate work way college her world open one professor challenge create culture service workplace put five component legendary service practice although ferguson store kelsey work certainly know service excellence kelsey believe make positive difference she quickly learn culture change easy role frontline employee significant ever could imagine in characteristic blanchard style legendary service the key be care quick entertain read people organizational level every industry when applied lesson profound impact service experience customer receive whether ceo part time employee every person make difference customer service everyone job
- 128 Label: business
- 129 Predicted class: ['business']
- 130
- 131 Sample description: from author 1 bestseller commercial real estate book amazon the due diligence handbook for

131 commercial real estate quit run investment property influence hopeium delusional state believe everything work fine without put forth effort think create value that way investor approach investment property management plan nbsp without strategy increase value property management plan take lottery approach nbsp you want cliff note version in word 20 give 80 result bunch minutia 30 50 page fluff sift find helpful tip key point if read book the due diligence handbook for commercial real estate know i talk i attribute popularity fact cut meat subject give reader step step instructional information need conduct due diligence properly buy investment property this extremely important part process make break investment decision your golden opportunity may turn lead know find hidden problem undiscovered issue value creation opportunity investment may hold much money save make learn properly unfortunately investor real estate professional know properly conduct due diligence know know nbsp the fundamental remain part come add value lease this information use across various genre commercial real estate investment whether office industrial retail multi family residential property i comment time time throughout book different angle strategy may want consider different property type i obtain background experience 30 year commercial real estate industry i know best particular specialty in handbook learn how determine best way price rental rate investment property get lease high value nbsp find way make property attractive tenant how interview choose right people help get property lease how structure lease rate term create value investment property how incentivize tenant get property lease crowded market strategy tip create value essential analysis form proposal lease form sample tenant application formstthese learnable skill put use right away enhance effort add value real estate investment for seasoned commercial real estate investor professionals i share valuable experience time landlord representative i lease property local lease broker i also spend time vice president lease investor own 12 million square foot across u s i learn local practice custom well ins outs various market los angeles dallas houston chicago orlando phoenix experience learn phase career also contribute content book by learn implement essential value enhance principle easily increase property cash flow exponentially every day wait another day flush money drain buy today start generate income instead hop work you look back consider one best real estate investment

132 Label: business

133 Predicted class: ['business']

134

```

135 ======TRAINING SGD=====
136 /Users/MikiWong/anaconda3/envs/scikitlearn/lib/python3.5/
site-packages/sklearn/linear_model/stochastic_gradient.py:
128: FutureWarning: max_iter and tol parameters have been
added in <class 'sklearn.linear_model.stochastic_gradient.
SGDClassifier'> in 0.19. If both are left unset, they
default to max_iter=5 and tol=None. If tol is not None,
max_iter defaults to max_iter=1000. From 0.21, default
max_iter will be 1000, and default tol will be 1e-3.
137 "and default tol will be 1e-3." % type(self),
FutureWarning)
138             precision    recall   f1-score   support
139
140 business      0.96823   0.97905   0.97361    716
141 cook          0.98879   0.97601   0.98236    542
142 literature    0.96084   0.98305   0.97182    649
143 sports         0.99334   0.96759   0.98030    617
144
145 avg / total   0.97688   0.97662   0.97666    2524
146
147 InterAnnotator Agreement score: 0.968718208227
148 ======TRAINING NEAREST NEIGHBOUR=====
149             precision    recall   f1-score   support
150
151 business      0.92627   0.96508   0.94528    716
152 cook          0.90986   0.98708   0.94690    542
153 literature    0.97569   0.86595   0.91755    649
154 sports         0.95765   0.95300   0.95532    617
155
156 avg / total   0.94313   0.94136   0.94095    2524
157
158 InterAnnotator Agreement score: 0.921606385787
159 ======TRAINING DECISION TREE=====
160             precision    recall   f1-score   support
161
162 business      0.86283   0.87849   0.87059    716
163 cook          0.94424   0.93727   0.94074    542
164 literature    0.89441   0.88752   0.89095    649
165 sports         0.92007   0.91410   0.91707    617
166
167 avg / total   0.90242   0.90214   0.90225    2524
168
169 InterAnnotator Agreement score: 0.869050812262
170 ======TRAINING RANDOM FOREST=====
171             precision    recall   f1-score   support
172
173 business      0.88542   0.94972   0.91644    716
174 cook          0.96998   0.95387   0.96186    542
175 literature    0.91871   0.92296   0.92083    649
176 sports         0.95797   0.88655   0.92088    617

```

```
177
178 avg / total 0.92987 0.92829 0.92841 2524
179
180 InterAnnotator Agreement score: 0.903950657931
181 ****
182 *****ENSEMBLE METHODS*****
183 =====TRAINING RANDOM FOREST=====
184             precision    recall   f1-score   support
185
186 business 0.89947 0.94972 0.92391 716
187 cook      0.95701 0.94465 0.95079 542
188 literature 0.93041 0.94761 0.93893 649
189 sports    0.96329 0.89303 0.92683 617
190
191 avg / total 0.93538 0.93423 0.93426 2524
192
193 InterAnnotator Agreement score: 0.911929990287
194 =====TRAINING ADABOOST=====
195             precision    recall   f1-score   support
196
197 business 0.88119 0.87011 0.87561 716
198 cook      0.97872 0.93358 0.95562 542
199 literature 0.82657 0.91063 0.86657 649
200 sports    0.96410 0.91410 0.93844 617
201
202 avg / total 0.90836 0.90491 0.90583 2524
203
204 InterAnnotator Agreement score: 0.872717301546
205
206 Process finished with exit code 0
207
```