

Όνομα	Επίθετο	A.M
Ένρι	Γκάτση	1115200900048
Gerald	Mema	1115200800108

Περίληψη

Στην αναφορά αυτήν θα περιγράψουμε με ποιον τρόπο δουλέψαμε πάνω στην εργασία και θα αναλύσουμε τα ζητούμενα της άσκησης και με ποιον τρόπο αυτά επιλύθηκαν και προσεγγίστηκαν. Αρχικά, να αναφέρουμε ότι φτιάχτηκαν 5 wordcloud , ένα για κάθε κατηγορία, τα οποία μας δίνουν μία εικόνα για την συχνότητα των λέξεων στα document στα οποία . Το πρώτο πράγμα που αναλύουμε είναι το preprocessing που κάναμε ή που σκεφτήκαμε να κάνουμε πριν εφαρμόσουμε τους αλγορίθμους κατηγοριοποίησης και τον αλγόριθμο k-means. Στην συνέχεια περιγράφουμε την υλοποίηση του k-means όπως έχει ζητηθεί από την εκφώνηση της άσκησης, ποιες τεχνικές χρησιμοποιήσαμε για την βελτιστοποίηση αυτού και τις παρατηρήσεις μας πάνω σε αυτές. Το επόμενο βήμα αναφέρεται στην κατηγοριοποίηση των δεδομένων, δηλαδή πως χρησιμοποιήθηκαν οι 4 αλγόριθμοι, τι αποτελέσματα έχει ο καθένας χωρίς καμία επεξεργασία και ποια απόδοση έχουν μετά, τα οποία σχολιάζονται και γίνονται παρατηρήσεις πάνω σε αυτά. Να επισημάνουμε εδώ πως στα αποτελέσματα αρχικά αναφέρουμε τι accuracy έχει ο κάθε αλγόριθμος και τι accuracy με 10-fold cross validation και στην συνέχεια το accuracy που έχει μετά την προ επεξεργασία των δεδομένων στις οποίες φυσικά και περιγράφουμε για ποιον λόγο έγιναν. Τέλος για τον δικό μας κατηγοριοποιητή αποφασίσαμε να φτιάξουμε έναν συνδυασμό από από τους classifier που είχαμε να μελετήσουμε εφόσον κάναμε αρκετές δοκιμές πάνω σε αυτούς.

Παρατηρήσεις

Παρατήρηση1: Στην προσπάθεια μας να βρούμε βέλτιστες λύσεις των παραμέτρων για καλύτερα αποτελέσματα από τους αλγορίθμους δεν μπορέσαμε να χρησιμοποιήσουμε GridSearch διότι δουλέψαμε την εργασία σε windows και έχει ένα πρόβλημα η συνάρτηση διότι ζητάει να κάνει fork και τα windows δεν το υποστηρίζουν οπότε χρειάζεται να βάλει κανείς if __name__ == __main__ στα κομμάτια που θέλει να εκτελεστούν μόνο από την κύρια διεργασία και γινόταν πολύ περίπολοκη η υλοποίηση, πράγμα το οποίο πιστεύουμε ότι δεν ήταν ο σκοπός της άσκησης.

Παρατήρηση2: Στην προσπάθεια προ-επεξεργασίας των δεδομένων χρησιμοποιήσαμε μία έτοιμη βιβλιοθήκη(Portstemmer) για να κάνει stem τις λέξεις, η οποία αυτή τεχνική δεν βοήθησε στην βελτίωση της απόδοσης και έτσι δεν εφαρμόστηκε.

Παρατήρηση3: Θεωρήσαμε πρέπον στην αναφορά αυτή να μην δώσουμε έμφαση στο τι αποτελέσματα μας δίνει κάθε μία δοκιμή αλλά στον σχολιασμό αυτών και στο γιατί βγάζουμε αυτά τα αποτελέσματα.

Παρατήρηση4: Για την αξιολόγηση των αλγορίθμων εστιάσαμε και προσπαθήσαμε να αυξήσουμε την απόδοση τους με βάση τα αποτελέσματα που έδινε το accuracy από 10-fold cross validation.

Παρατήρηση5: Σύμφωνα με τα ζητούμενα της εργασίας δώσαμε ιδιαίτερο βάρος στην προεπεξεργασία των δεδομένων και στις ιδέες για πως θα ήταν καλό να επεξεργαστούν τα δεδομένα για να δουλέψει

πιο αποδοτικά ο αλγόριθμος.

Παρατήρηση6: Στα αρχεία κώδικα που θα παραδώσουμε οι τιμές των παραμέτρων που έχουμε αφήσει είναι αυτές που δοκιμάστηκαν κατά την τελευταία δοκιμή. Έχουμε αφήσει τις καλύτερες τιμές (αυτές που αναλύσαμε παραπάνω) , στις οποίες μπορούμε να δοκιμάσουμε καινούργια δεδομένα στο MyMethod.py

Αρχεία

Για να τρέξουν τα προγράμματα χρειάζονται τα αρχεία test_set και train_set.

10_fold_tester.py: Καλεί τους classifiers οι οποίοι κάνουν μόνοι τους cross-validation και παράγει το αρχείο EvaluationMetric_10fold.csv

k-means.py: Υλοποίηση του k-means.

MyMethod.py: Υλοποίηση της δικιάς μας μεθόδου

MyMethodForTestSet.py: Εκτελεί τη δικιά μας μέθοδο και παράγει το αρχείο testSet_categories

Τα υπόλοιπα αρχεία εκτελούν τους classifiers με τις βέλτιστες παραμέτρους που βρήκαμε

Preprocessing

Τα βήματα που ακολουθήθηκαν για την προ επεξεργασία των δεδομένων είναι τα εξής:

1)Cleaning: Στο βήμα αυτό έγινε καθαρισμός των αρχικών δεδομένων. Πιο αναλυτικά χρησιμοποιήθηκε η βιβλιοθήκη stopwords για να πετάξουμε τις λέξεις που δεν θα συμβάλουν στην ανάλυση και κατηγοριοποίηση των δεδομένων. Η διαδικασία αυτή έγινε με διάφορους τρόπους, όπως να χρησιμοποιηθεί το wordcloud που φτιάχτηκε αρχικά για κάθε κατηγορία του train_set αλλά κυρίως χρησιμοποιήθηκε η βοήθεια από τους vectorizers με τους οποίους μπορούσαμε να αγνοήσουμε λέξεις ανάλογα με την βαρύτητά τους(παραμέτροι max_df, min_df). Η ιδέα είναι να πετάξουμε τις πιο συχνές εμφανιζόμενες λέξεις με το σκεπτικό ότι δεν συνεισφέρουν στην απόφαση για το αν ένα κείμενο ανήκει σε μια συγκεκριμένη κατηγορία ή όχι. Επίσης ούτε η πιο σπάνιες λέξεις δεν θα μπορούσαμε να πούμε ότι ανήκουν στις λέξεις κλειδιά διότι δεν βοηθάνε στον υπολογισμό ομοιοτήτων. Επομένως κρατάμε κάτι ενδιάμεσο σαν γενική ιδέα το οποίο βέβαια αλλάζει ανάλογα με τον ποιο classifier χρησιμοποιούμε κάθε φορά. Η μέθοδος του cleaning εφαρμόζεται και στον αλγόριθμο k-means για clustering.

2)Format: Πέρα από το να πετάμε λέξεις από το κείμενο είναι σημαντικό να διαμορφώνουμε τα δεδομένα μας έτσι ώστε να έχουμε καλύτερη εικόνα και να μπορεί να γίνει καλύτερη ανάλυση πάνω σε αυτά. Για τον λόγο αυτό χρησιμοποιήθηκε η βιβλιοθήκη stem έτσι ώστε λέξεις που μοιάζουν μεταξύ τους, (π.χ steps, step) να τις αποτυπώνουμε ως μία ενιαία λέξη. Η ιδέα αυτή έχει μία λογική αφού θα η ομοιότητα μεταξύ 2 κειμένων θα υπολογίζεται πιο αποτελεσματικά. Η μέθοδος αυτή σίγουρα έχει και ρίσκο αφού μπορεί να ενώσει 2 λέξεις που έχουν διαφορετική σημασία. Στην εργασία η μέθοδος αυτή δεν βελτίωσε ιδιαίτερα τους αλγόριθμους και σε κάποιες φορές η απόδοση έπεφτε.

3)Normalization: Για να εκτελέσουμε τους αλγόριθμους που έχουμε να εξετάσουμε θα πρέπει να μετατρέψουμε τα κείμενα σε διανύσματα τα οποία διανύσματα μας δίνουν το πλήθος των φορών που εμφανίζεται μία λέξη και τι βάρος θα έχει αυτή. Όμως όπως γνωρίζουμε οι αλγόριθμοι υπολογίζουν αποστάσεις ,άρα τα δεδομένα μας θα πρέπει να ομαλοποιηθούν σε ένα καινούργιο διάνυσμα το οποίο

θα έχει συνολικό άθροισμα ίση με την μονάδα και αυτό το επιτυγχάνουμε με την χρήση του TfIdfTransformer. Ένα επίσης σημαντικό κομμάτι της ομαλοποίησης των δεδομένων είναι και το dimensionality reduction, όπου είναι η διαδικασία με την οποία μειώνουμε τις διαστάσεις του προβλήματος και προβάλλουμε τις σημαντικότερες από αυτές. Αυτό επιτυγχάνεται με την χρήση του αλγόριθμου lsi και χρησιμοποιήσαμε την βιβλιοθήκη TruncatedSVD. Παρακάτω αναλύουμε τις διάφορες δοκιμές που έχουμε κάνει με βάση το πλήθος των διαστάσεων που κρατάμε για κάθε αλγόριθμο.

4)Χρησιμοποίηση του τίτλου για την πιο αποτελεσματική κατηγοριοποίηση των δεδομένων:

Αυτό που σκεφτήκαμε με τον τίτλο είναι να τον πάρουμε όπως έχει δηλαδή σαν κείμενο και να το βάλουμε πολλές φορές μέσα στο περιεχόμενο, δίνοντάς του έτσι βάρος από πλευράς CountVectorizer. Το θέμα είναι ότι αυτό δεν ήταν και τόσο επιτυχημένο όταν εφαρμοζόταν cross validation. Για παράδειγμα στον naive-bayes όταν βάζαμε τον τίτλο 70 φορές στο κείμενο τότε φτάναμε ακόμα και 99% accuracy και f-measure αλλά στα καινούργια δεδομένα (test data in cross validation) έπεφτε στο 93%. Αυτό δηλαδή που συμπεράνουμε είναι ότι ο τίτλος να μην είναι σημαντικός στην κατηγοριοποίηση αλλά υπάρχει κίνδυνος overfitting.

Clustering

Περιγραφή αλγορίθμου k-means

Συνοπτικά θα περιγράψουμε τον αλγόριθμο αυτό και τα βήματά του:

1. Βρες τα αρχικά 5 κέντρα
 - α)πάρε αυτά τυχαία
 - β)υλοποίηση αλγορίθμου k-means++, έτσι ώστε να 5 αρχικά κέντρα να μην είναι τυχαία αλλά βρίσκονται με κάποια πιθανότητα πιο μακριά μεταξύ τους
2. Όσο τα κέντρα που βρίσκεις από δω και πέρα είναι διαφορετικά από τα αμέσως προηγούμενα κέντρα τότε
 - α)υπολόγισε αποστάσεις σημείων με τα κέντρα τους(cosine similarity) και βάλε το σημείο στο cluster με το μικρότερη απόσταση
 - β)υπολόγισε ξανά τα κέντρα μεταξύ των νέων σημείων που μπήκαν στα cluster.
 - γ)Επανέλαβε το Βήμα 2

Επίσης ένας περιορισμός που λαμβάνεται υπόψη είναι ο αλγόριθμος να μην ξεπερνάει τις 100 επαναλήψεις.

Δοκιμές

Κάθε φορά που εκτελούμε τον αλγόριθμο τα αποτελέσματα γράφονται σε ένα results.csv όπως ζητήται από την εκφώνηση της άσκησης. Τα αποτελέσματα είναι ικανοποιητικά σε μεγάλο βαθμό. Με preprocessing να διώχνουμε συχνά εμφανιζόμενες λέξεις και πολύ σπάνιες λέξεις (χρησιμοποιήθηκαν τα max_df, min_df στους vectorizers). Βέβαια επειδή ο αλγόριθμος εξαρτάται πολύ από το βήμα 1 , δηλαδή πως θα επιλέξει τα αρχικά κέντρα χρειάστηκε να επαναλάβουμε πολλές φορές την διαδικασία. Ένα παράδειγμα εκτέλεσης του αλγορίθμου με CountVectorizer:

1.CounterVector=CountVectorizer(stop_words=stopWords,max_df=0.55,min_df=0.005)

δηλαδή διώχνοντας κάποιες λέξεις από το wordcloud και τις λέξεις που εμφανίζονται περίπου μια φορά στα 2 κείμενα αλλά και αυτές που εμφανίζονται λιγότερες από μια φορά στα 200 κείμενα .

2.FrequencyCounterVector=transformer.fit_transform(CounterVector)

Ομαλοποίηση του διανύσματος CounterVector που μετράει πόσες φορές υπάρχει η λέξη στο κείμενο.

Σε αυτήν την περίπτωση η ομαλοποίηση , όταν χρησιμοποιήθηκε η Ευκλείδεια απόσταση,είχε μεγάλη

σημασία αφού το π.χ το 126 με το 100 έχει μεγαλύτερη απόσταση από το 20 με το 0 το οποίο θα μας δώσει λάθος αποτελέσματα όσον αφορά την ομοιότητα μεταξύ αυτών των σημείων, αλλά στην περίπτωση του cosine similarity δεν έχει και τόση μεγάλη διαφορά αφού υπολογίζουμε και κρατάμε την μικρότερη γωνία αλλά και πάλι κρατάμε την ομαλοποίηση αφού μας δίνει καλύτερα αποτελέσματα.

3.svd=TruncatedSVD(n_components=100, random_state=42)

Σε αυτό το βήμα κάνουμε dimensionality reduction και μειώνουμε το πλήθος των διαστάσεων. Στην συγκεκριμένη περίπτωση το 100 είναι μια καλή τιμή συνδυασμού χρόνου και απόδοσης.

Τα αποτελέσματα είναι:

Cluster1:	business = 0.928 = 2537/2735
	film=0.9875 = 2212/2240
	football=0.003 = 11/3121
	politics=0.014 = 37/2683
	technology=0.031 = 47/1487
Cluster2:	business = 0.005 = 15/2735
	film=0.003 = 8/2240
	football=0.004 = 12/3121
	politics=0.096 = 257/2683
	technology=0.02 = 30/1487
Cluster3:	business = 0.047 = 128/2735
	film=0.007 = 15/2240
	football=0.0006 = 2/3121
	politics=0.0096 = 26/2683
	technology=0.933 = 1388/1487
Cluster4:	business = 0.003 = 8/2735
	film=0.001 = 3/2240
	football=0.988 = 3084/3121
	politics=0.000 = 0/2683
	technology=0.008 = 12/1487
Cluster5:	business = 0.017 = 47/2735
	film=0.0009 = 2/2240
	football=0.004 = 12/3121
	politics=0.881 = 2363/2683
	technology=0.007 = 10/1487

Αντίστοιχο παράδειγμα έγινε και με την χρήση του TfidfVectorizer και παρατηρήθηκε ότι στην περίπτωση αυτή επειδή κρατάμε τα βάρη στις λέξεις μπορούμε να διώξουμε παραπάνω λέξεις, δηλαδή μπορούμε να βάλουμε την τιμή της μεταβλητής max_df=0.2 που σημαίνει ότι αν μια λέξη εμφανίζεται παραπάνω από μια φορά στα 5 κείμενα, αγνοήσε την και έτσι θα περιορίσουμε ακόμα πιο πολύ το dataset που έχουμε να εξετάσουμε και θα εμβαθύνουμε πιο πολύ στις σημαντικότερες λέξεις του κειμένου.

Σχολιασμός

Για την χρονική επίδοση του k-means φτιάξαμε τον k-means++ όπως αναφέραμε προηγουμένως έτσι ώστε να μειώσουμε τον αριθμό των επαναλήψεων του αλγορίθμου και τις περισσότερες φορές αυτό επιτυγχάνεται. Αυτό όμως είχε σαν συνέπεια να καθυστερούμε παραπάνω στον υπολογισμό των αρχικών 5 κέντρων και επίσης δεν έχουμε καμία εγγύηση ότι δεν θα έχουμε πολλές επαναλήψεις διότι τα κέντρα μπορεί να μην συγκλίνουν αμέσως. Έχουμε κρατήσει την περίπτωση του k-means++ και η τυχαία μέθοδος είναι σχολιασμένη στον κώδικα.

Επίσης παρατηρήθηκε αρκετές φορές το φαινόμενο ότι μπορεί μία κατηγορία να την “έσπαγε” σε 2 cluster και σε άλλο cluster να έβαζε 2 κατηγορίες μαζί. Επίσης παρατηρήθηκε ότι η 1η περίπτωση αφορούσε τις πιο πολλές φορές την κατηγορία football και η δεύτερη την κατηγορία business και technology. Για να μειώσουμε το παραπάνω φαινόμενο σκεφτήκαμε να βάζουμε στο περιεχόμενο της κατηγορίας football τον τίτλο του, η μάλλον τις λέξεις του τίτλου πολλές φορές έτσι ώστε να αυξήσουμε το βάρος τους με το CountVectorizer.

Γενικά αυτά που βοήθησαν πάρα πολύ είναι οι παράμετροι max_df, min_df που δίνουμε στους vectorizers οι οποίοι αγνοούσαν λέξεις οι οποίες θα μπορούσαν να βγάλουν ψευδείς ομοιότητες μεταξύ των κειμένων.

Σημείωση: Αν εκτελεστεί το kmeans.py παράγει σαν έξοδο ένα csv αρχείο με όνομα clustering_Kmeans.csv σύμφωνα με τις προδιαγραφές της εργασίας.

Classification

Σε αυτό το κομμάτι της εργασίας μπορούμε να μιλήσουμε πιο συγκεκριμένα για την απόδοση των πειραμάτων μας αφού ο k-means ακολουθεί μια μη ντετερμινιστική προσέγγιση του προβλήματος. Παρακάτω θα αναφέρουμε και θα σχολιάσουμε τους 4 αλγόριθμους που έχουμε για εξέταση και τα πειράματα που κάναμε πάνω σε αυτά.

1)Naive-bayes

Ο Naive Bayes είναι ένας probabilistic αλγόριθμος ο οποίος βήμα του preprocessing δεν μπορούμε να κάνουμε dimensionality reduction με την χρήση του lsi και ο λόγος είναι ότι η lsi μπορεί να βγάλει αρνητικές τιμές κάτι που ο Naive Bayes δεν μπορεί να υπολογίσει διότι δεν μπορεί να υπολογιστεί η πιθανότητα όταν έχουμε μία λέξη η οποία εμφανίζεται αρνητικές φορές στο κείμενο. Για την εργασία αυτή έχουμε χρησιμοποιήσει το MultinomialNB.

Train data accuracy/test data accuracy(average10foldcross validation) without preprocessing

accuracy with CountVectorizer = **0.97**

accuracy with CountVectorizer and TfidfTransformer= **0.93**

accuracy with TfidfVectorizer= **0.96**

accuracy(10-fold cross validation) with CountVectorizer = **0.96**

accuracy(10-fold cross validation) with CountVectorizer and TfidfTransformer=**0.91**

accuracy(10-fold cross validation) with TfidfVectorizer = **0.94**

Περιγραφή δοκιμών , σχολιασμός και παρατηρήσεις:

Σε αυτόν τον αλγόριθμο θα εστιάσουμε περισσότερο στο CountVectorizer αφού όπως παρατηρούμε έχουμε μια πολύ καλή απόδοση και όταν μας έρχονται καινούργια δεδομένα.

Οι δοκιμές που κάναμε ήταν να χρησιμοποιήσουμε τον τίτλο τοποθετώντας(αρχικά) τις λέξεις του 100 φορές μέσα στο περιεχόμενο. Αυτό είχε σαν συνέπεια να φτάσουμε το accuracy 99% αλλά όταν εφαρμόζαμε cross validation το accuracy μειώνεται στο 93% , αυτό σημαίνει ότι πρέπει κάπως

να ομαλοποιήσουμε αυτήν την διαφορά αφού παρατηρείτε overfitting, δηλαδή ο αλγόριθμος λειτουργεί τέλεια μόνο για τα δεδομένα που κάνει fit. Στην συνέχεια προσπαθήσαμε να δώσουμε έμφαση έτσι ώστε να 'πετάξουμε' πολύ εμφανιζόμενες λέξεις με την χρήση της μεταβλητής max_df του vectorizer. Μετά από αρκετές δοκιμές καταλήξαμε ότι αν αγνοήσεις τις λέξεις που εμφανίζονται πάνω από το 60% στο πλήθος των κειμένων και τοποθετήσεις τις λέξεις του τίτλου 10 φορές στο περιεχόμενο τότε παίρνεις accuracy 98% και accuracy με cross validation 97%.

Η παρατήρησή μας για τα αποτελέσματα είναι ότι ο τίτλος είναι χρήσιμο εργαλείο για τον αλγόριθμο αυτό και η σκέψη μας ήταν να δώσουμε βάρος σε αυτό βάζοντας πολλές φορές αυτές τις λέξεις μέσα στο περιεχόμενο. Μία ακόμα αξιοσημείωτη παρατήρηση που πρέπει να κάνουμε είναι ότι ο αλγόριθμος αυτός βγάζει πολύ καλά αποτελέσματα(κοντά στο 98%) ακόμα και αν αγνοήσουμε πάρα πολλές λέξεις (max_df = 0.002). Αυτό πιστεύουμε οφείλεται στο ότι εδώ δεν μετράμε αποστάσεις αλλά υπολογίζεται πιθανότητα, οπότε δεν χρειαζόμαστε και τόσο μεγάλο όγκο δεδομένων για να πάρουμε μία καλή απόφαση αρκεί να έχουμε κάποιες όμοιες λέξεις. Τέλος αυτό που είναι σημαντικό να πούμε είναι ότι σε σχέση με τους άλλους αλγόριθμους στον Naive Bayes δεν χρειάζεται ομαλοποίηση το διάνυσμα που κρατάει το πλήθος εμφάνισης των λέξεων, αντιθέτως βγάζει χειρότερα αποτελέσματα. Και σε αυτήν την περίπτωση ισχυριζόμαστε ότι συμβαίνει λόγω πιθανότητας που υπολογίζει, δηλαδή είναι πιο χρήσιμο να γνωρίζει το πλήθος των φορών που εμφανίζεται μία λέξη.

2)K-Nearest Neighbors

Ο αλγόριθμος αυτός κατηγοριοποιεί τα κείμενα με βάση τους κοντινότερους γείτονες ,το πλήθος των οποίων δίνουμε σαν παράμετρο στην συνάρτηση αυτού του classifier. Τα βήματα που ακολουθούμε είναι αρχικά μετατροπή κειμένου σε διάνυσμα και μετά μείωση των διαστάσεων του προβλήματος.

Train data accuracy/test data accuracy(10-fold cross validation) without preprocessing(components=50,n_neighbors=5)

accuracy with CountVectorizer = **0.88**

accuracy with CountVectorizer and TfidfTransformer= **0.96**

accuracy with TfidfVectorizer= **0.96**

accuracy(10-fold cross validation) with CountVectorizer = **0.81**

accuracy(10-fold cross validation) with CountVectorizer and TfidfTransformer=**0.95**

accuracy(10-fold cross validation) with TfidfVectorizer = **0.94**

Περιγραφή δοκιμών , σχολιασμός και παρατηρήσεις:

Οι δοκιμές που κάναμε και σε αυτήν την περίπτωση είναι βασισμένη στην προ επεξεργασία που αναφέρουμε παραπάνω. Αρχικά κάναμε την δοκιμή με k=1 και τα αποτελέσματα είναι 100% επιτυχία χωρίς cross validation αλλά δεν έχουμε το ίδιο ποσοστό ευστοχίας όταν έρχονται νέα δεδομένα αλλά και πάλι είναι αρκετά ικανοποιητικά(96%). Εμείς θα προσπαθήσουμε να βελτιώσουμε τα αρχικά αποτελέσματα για k=5.

Σε αυτόν τον αλγόριθμο εστίασαμε στο να κάνουμε αρχικά CountVectorizer και μετά TfidfTransformer αφού η ομαλοποίηση του διανύσματος είναι απαραίτητη για τον καλύτερο υπολογισμό της ομοιότητας(απόστασης). Επίσης χρησιμοποιούμε τον τίτλο για να αυξήσουμε την απόδοσή μας τοποθετώντας τον 10 φορές μέσα στο κείμενο και επίσης θα ήταν καλό να διώξουμε και τις πιο συχνές λέξεις (max_df=0.85), έτσι το accuracy αυξάνεται κατά 1% (97% και 96% για cross validation). Αν αυξήσουμε τον αριθμό των διαστάσεων από 50 σε 100 τότε έχουμε άλλο 1% αύξηση.

Σε αντίθεση με τον Naive Bayes παρατηρούμε ότι σε αυτήν την περίπτωση χρειαζόμαστε αρκετή πληροφορία(λέξεις) για να έχουμε καλύτερα αποτελέσματα και ο λόγος που πιστεύουμε ότι γίνεται αυτό είναι ότι μπορεί να πεταχτεί μία σημαντική λέξη η οποία να καθόριζε το τελικό αποτέλεσμα. Επίσης παρατηρήθηκε ότι όσο μικρότερο k έχουμε τόσο καλύτερα αποτελέσματα βγάζουμε και είναι απολύτως λογικό αφού ψάχνουμε πιο συγκεκριμένα. Να αναφέρουμε τέλος ότι ρόλο απόδοση έπαιξε και ο αριθμός των διαστάσεων(δοκιμάστηκε μέχρι 200) όπου όσο μεγαλύτερος τόσο πιο ικανοποιητικά ήταν τα αποτελέσματα.

3)Random Forest

Με τον αλγόριθμο Random Forest χωρίς να κάνουμε 10-fold cross validation και χωρίς καμία προ επεξεργασία παρατηρούμε ότι η ακρίβεια είναι περίπου ίσο με 0,997. Δηλαδή το κάνουμε train με το df['Content'] και ύστερα κάνουμε predict πάνω στο df['Content'] και είμαστε σχεδόν πάντα μέσα στις προβλέψεις μας. Αυτό αιτιολογείται λόγω του γεγονότος πως ο αλγόριθμος για τα δεδομένα τα οποία έχει για να κάνει train μπορεί να χωρίσει το χώρο σε κατάλληλα κομμάτια ώστε να παίρνει πολύ καλές αποφάσεις. Αυτό όμως αποτελεί πρόβλημα όταν πάμε να κάνουμε cross-validation διότι επειδή ακριβώς ο χώρος των δεδομένων έχει χωριστεί υπερβολικά καλά ο αλγόριθμος ταξινομεί νέα δεδομένα σε λάθος κλάσεις και παρατηρούμε τη δραματική πτώση της ακρίβειας στο 0.94, γνωστό φαινόμενο σαν overfitting. Το ενδιαφέρον μας εστιάζει στα αποτελέσματα του cross-validation καθότι μας ενδιαφέρει να έχουμε έναν κατηγοριοποιητή ο οποίος μπορεί να κατηγοριοποιεί ορθά νέα δεδομένα εισαγωγής και γι' αυτό ακολουθούν τα εξής πειράματα με σκοπό τη πορεία σε αυτή τη κατεύθυνση.

Train data accuracy/test data accuracy(10-fold cross validation) without preprocessing

accuracy with CountVectorizer= **1.00**

accuracy with TfidfVectorizer and TfidfTransformer= **1.00**

accuracy with TfidfVectorizer= **1.00**

accuracy(10-fold cross validation) with CountVectorizer = **0.78**

accuracy(10-fold cross validation) with CountVectorizer and TfidfTransformer=**0.94**

accuracy(10-fold cross validation) with TfidfVectorizer = **0.94**

Περιγραφή δοκιμών , σχολιασμός και παρατηρήσεις:

Αλλάζοντας το πλήθος των δέντρων παρατηρούμε βελτίωση των αποτελεσμάτων και άνοδο της ακρίβειας γύρω στο 96%. Αυτό εξηγείται από το γεγονός πως η κατηγοριοποίηση έχει περισσότερα δέντρα να συμβουλευτεί οπότε μπορεί να αποκτήσει περισσότερες σωστές “γνώμες” για να παράγει ο αλγόριθμος ένα τελικό αποτέλεσμα.

Δοκιμάσαμε να τρέξουμε τον αλγόριθμο με τη μεταβλητή warm start = TRUE και τα αποτελέσματα ήταν πολύ ικανοποιητικά αφού με cross validation από την 2η φορά και μετά φτάνει στο 1.0. Εν κατακλείδι οι υπόλοιπες δοκιμές που κάναμε πειράζοντας τα ορίσματα της συνάρτησης δεν παρουσίασαν κάποια ιδιαίτερη βελτίωση στα αποτελέσματα.

4)Support Vector Machine

Ίσως ο πιο αργός αλγόριθμος από τους 3 προηγούμενους και ο πιο αποτελεσματικός. Υπάρχουν πολλοί μέθοδοι με τους οποίους μπορούμε να τον προσεγγίσουμε αλλά εμείς έχουμε επιλέξει τον πιο απλό που είναι το γραμμικό svm. Η αποτελεσματικότητά του οφείλεται κυρίως στο ότι δοκιμάζει αρκετές εναλλακτικές λύσεις και αποφασίζει ποιος για την καλύτερη από αυτήν. Η απόφασή του έχει λογικό περιεχόμενο αφού αποφασίζει αν ένα σημείο ανήκει σε μια κατηγορία ή όχι.

Train data accuracy/test data accuracy(10-fold cross validation) without preprocessing

accuracy with CountVectorizer = **0.90**

accuracy with CountVectorizer and TfidfTransformer= **0.96**

accuracy with TfidfVectorizer= **0.96**

accuracy(10-fold cross validation) with CountVectorizer = -

accuracy(10-fold cross validation) with CountVectorizer and TfidfTransformer=**0.96**

accuracy(10-fold cross validation) with TfidfVectorizer = **0.96**

Περιγραφή δοκιμών , σχολιασμός και παρατηρήσεις:

Με μία πρώτη ματιά μπορούμε να συμπεράνουμε την απόδοση του αλγορίθμου αυτού παρατηρώντας ότι με και χωρίς cross validation το αποτέλεσμα. Αυτό οφείλεται κυρίως στο ότι ο αλγόριθμος αυτός εμπεριέχει 5-fold cross validation μέσα στην υλοποίηση του. Αυτό αποφεύγει το overfitting και μπορούμε απλά με το κατάλληλη προ επεξεργασία να βελτιώσουμε το accuracy του.

Στην μέθοδο αυτή κάναμε την ίδια διαδικασία προ-επεξεργασίας με τους προηγούμενους αλγόριθμους, χρησιμοποιώντας τον τίτλο και την αποκοπή πολύ χρησιμοποιημένων λέξεων αυξάνοντας έτσι την απόδοση του αλγορίθμου(accuracy=0.97). Παρατηρήθηκε επίσης ότι μπορούμε να έχουμε καλά αποτελέσματα κρατώντας έναν μικρό όγκο δεδομένων(max_df=0.1).

MyMethod

Στο ερώτημα αυτό χρησιμοποιήθηκε ένας συνδυασμός από τους 4 αλγορίθμους που είχαμε να εξετάσουμε. Η επιλογή αυτή έγινε διότι εφαρμόσαμε αρκετές δοκιμές πάνω στην προ-επεξεργασία των δεδομένων (διαφορετική για κάθε αλγόριθμο), και έτσι αποφασίσαμε να κρατήσουμε τα αποτελέσματα των classifier και με βάση της κατηγορίας που προέβλεψε ο καθένας εμείς κρατάμε την κατηγορία που προέβλεψε η πλειοψηφία.

Η μέθοδός μας ακολουθεί τα εξής βήματα.

- Χώρισε το dataset σε train και test
- Δώσε στους classifiers που είχαμε να ελέγξουμε το train_set και εκπαιδευσε τους.
- Πάρε τις προβλέψεις που έχουν υπολογίσει οι classifiers
- Βρες σε ποιά κατηγορία πιστεύουν οι περισσότεροι (classifiers) ότι ανήκει το έγγραφο του test_set και αποθήκευσε τον στο τελικό πίνακα των προβλέψεων
- Τελικά ελέγχουμε το accuracy

Τελικά βλέπουμε ότι παίρνουμε accuracy καλύτερο από όλους τους υπόλοιπους classifiers.

Πίνακας με αποτελέσματα του EvaluationMetric_10fold.csv

Statistic""Measure	Naive""Bayes	KNN	SVM	Random""Forest	My""Method
accuracy	0.967453213995	0.951179820993	0.961757526444	0.956875508544	0.964198535395
accuracy	0.968241042345	0.960912052117	0.960912052117	0.946254071661	0.96335504886
accuracy	0.954397394137	0.956026058632	0.951954397394	0.917752442997	0.958469055375
accuracy	0.964140179299	0.954360228199	0.952730236349	0.927465362673	0.960065199674
accuracy	0.960065199674	0.961695191524	0.962510187449	0.905460472698	0.965770171149
accuracy	0.957585644372	0.94942903752	0.951876019576	0.882544861338	0.954323001631
accuracy	0.941272430669	0.944535073409	0.942088091354	0.885807504078	0.945350734095
accuracy	0.954285714286	0.964897959184	0.953469387755	0.891428571429	0.961632653061
accuracy	0.948571428571	0.948571428571	0.943673469388	0.882448979592	0.947755102041
accuracy	0.951836734694	0.952653061224	0.943673469388	0.879183673469	0.951020408163

Πίνακας με αποτελέσματα του k-means

Business	Films	Football	Politics	Technology
0.003656307129798903	0.9875	0.003524511374559436	0.012299664554603056	0.03564223268325
0.014625228519195612	0.003125	0.0028836911246395386	0.8788669399925456	0.00672494956287
0.0032906764168190127	0.0008928571428	0.9913489266260814	0.0003727171077152441	0.00537995965030
0.8314442413162706	0.0017857142857	0.0019224607497596924	0.0980245993291092	0.00806993947545
0.1469835466179159	0.0066964285714	0.0003204101249599487	0.010436079016026835	0.94418291862811

Πίνακας με αποτελέσματα του testSet_categories.csv

ID	Predicted Category
2	Politics
10	Technology
25	Technology
28	Business
29	Business
33	Business
34	Business
37	Technology
39	Technology
40	Technology
44	Business
56	Business
66	Technology

75	Business
77	Business
85	Business
89	Business
92	Technology
107	Technology
112	Technology
115	Technology
139	Business
140	Business
152	Business
156	Technology
162	Technology
167	Business
170	Business
171	Technology
173	Politics
182	Technology
186	Football
193	Technology
196	Business
210	Business
218	Football
224	Business
227	Technology
230	Politics
232	Football
247	Football
262	Football
273	Football
276	Politics
279	Football
284	Business
286	Football
287	Business
305	Business
316	Business
337	Politics
342	Business
349	Film

357	Film
363	Film
367	Technology
368	Film
374	Technology
377	Film
385	Film
395	Film
401	Film
417	Film
422	Film
427	Film
433	Business
438	Business
469	Politics
470	Politics
473	Politics
481	Technology
486	Business
490	Business
495	Technology
510	Politics
519	Politics
523	Politics
533	Business
536	Politics
538	Politics
545	Business
550	Business
551	Business
557	Business
558	Business
561	Business
565	Business
575	Business
583	Business
589	Business
592	Business
593	Business
594	Business

598	Business
599	Politics
607	Business
614	Business
626	Business
627	Politics
634	Technology
636	Business
639	Business
641	Technology
643	Business
650	Business
655	Business
662	Technology
666	Technology
667	Business
673	Technology
678	Technology
693	Technology
694	Technology
696	Technology
697	Business
699	Business
704	Business
706	Business
707	Technology
709	Technology
717	Technology
724	Business
728	Technology
732	Technology
733	Business
741	Technology
746	Business
747	Business
749	Business
750	Politics
754	Business
756	Business
768	Politics

769	Business
774	Business
775	Business
776	Politics
789	Business
802	Politics
804	Politics
809	Technology
816	Technology
825	Technology
827	Film
830	Business
837	Technology
838	Business
839	Business
849	Business
854	Technology
865	Business
871	Politics
873	Politics
884	Politics
886	Politics
893	Business
896	Business
898	Business
907	Politics
909	Politics
910	Business
911	Politics
914	Politics
916	Politics
922	Business
925	Business
926	Politics
935	Business
936	Business
946	Politics
952	Business
954	Business
959	Business

960	Business
961	Business
964	Business
965	Business
968	Business
971	Business
972	Business
982	Business
983	Business
985	Business
987	Business
990	Politics
994	Business
1010	Politics
1014	Business
1024	Politics
1026	Business
1032	Business
1047	Politics
1050	Business
1059	Business
1061	Business
1062	Business
1069	Business
1072	Business
1074	Business
1081	Business
1088	Business
1090	Business
1093	Business
1097	Business
1100	Business
1102	Business
1105	Business
1142	Business
1143	Film
1155	Business
1162	Business
1163	Technology
1169	Business

1174	Business
1177	Technology
1178	Business
1184	Business
1185	Business
1192	Business
1208	Politics
1209	Politics
1215	Business
1217	Politics
1219	Business
1221	Business
1223	Business
1225	Politics
1229	Business
1230	Politics
1239	Politics
1251	Politics
1265	Technology
1269	Technology
1282	Football
1287	Technology
1295	Technology
1301	Technology
1302	Technology
1306	Film
1310	Technology
1314	Business
1322	Business
1328	Business
1331	Business
1332	Business
1337	Business
1345	Business
1349	Politics
1353	Politics
1356	Technology
1359	Technology
1368	Technology
1371	Business

1377	Technology
1381	Technology
1387	Technology
1389	Technology
1404	Technology
1411	Technology
1413	Technology
1425	Technology
1428	Business
1429	Business
1430	Technology
1433	Technology
1437	Technology
1438	Technology
1446	Technology
1447	Technology
1454	Business
1461	Business
1462	Business
1463	Technology
1466	Technology
1469	Technology
1475	Technology
1485	Technology
1488	Business
1493	Technology
1496	Film
1497	Film
1500	Technology
1502	Technology
1508	Business
1509	Technology
1518	Film
1522	Film
1524	Film
1527	Technology
1529	Film
1531	Film
1533	Film
1537	Film

1548	Film
1555	Film
1556	Film
1558	Film
1568	Film
1570	Film
1571	Film
1572	Film
1583	Film
1589	Film
1597	Film
1602	Film
1606	Film
1607	Film
1613	Film
1624	Film
1637	Film
1638	Film
1639	Film
1652	Film
1654	Film
1663	Politics
1675	Business
1695	Politics
1703	Politics
1704	Politics
1705	Politics
1709	Business
1714	Business
1715	Business
1725	Business
1726	Business
1741	Film
1749	Business
1755	Business
1760	Football
1761	Football
1768	Film
1772	Football
1775	Football

1778	Business
1779	Business
1781	Business
1784	Film
1789	Business
1797	Business
1803	Film
1805	Business
1812	Film
1819	Business
1833	Film
1843	Politics
1844	Politics
1845	Business
1860	Film
1864	Film
1873	Film
1878	Film
1879	Film
1884	Film
1885	Film
1889	Technology
1895	Film
1907	Film
1908	Film
1910	Film
1915	Film
1917	Film
1920	Film
1921	Film
1928	Film
1951	Film
1962	Film
1963	Technology
1964	Film
1965	Technology
1967	Film
1968	Film
1972	Film
1981	Film

1985	Film
1987	Film
1988	Film
1990	Film
1991	Film
1992	Film
1996	Film
2004	Film
2007	Film
2008	Film
2028	Film
2037	Film
2063	Film
2076	Film
2079	Football
2081	Technology
2083	Technology
2086	Technology
2100	Politics
2101	Politics
2103	Politics
2104	Film
2116	Film
2117	Politics
2118	Politics
2119	Film
2132	Film
2134	Film
2136	Film
2138	Film
2151	Politics
2154	Film
2158	Film
2164	Film
2168	Politics
2172	Film
2181	Film
2191	Football
2192	Film
2193	Film

2198	Film
2199	Film
2205	Film
2207	Film
2214	Film
2224	Film
2228	Film
2229	Film
2230	Film
2232	Politics
2244	Politics
2246	Film
2259	Politics
2266	Politics
2275	Politics
2277	Politics
2278	Politics
2279	Politics
2288	Politics
2300	Politics
2301	Politics
2307	Politics
2315	Politics
2318	Politics
2320	Politics
2323	Politics
2324	Politics
2325	Politics
2340	Politics
2341	Politics
2348	Politics
2350	Politics
2353	Politics
2362	Politics
2375	Politics
2377	Politics
2380	Politics
2383	Politics
2387	Politics
2391	Politics

2394	Politics
2395	Politics
2399	Politics
2410	Politics
2411	Politics
2415	Politics
2422	Politics
2424	Politics
2426	Politics
2432	Politics
2436	Politics
2437	Politics
2445	Politics
2447	Politics
2448	Politics
2464	Politics
2468	Politics
2473	Politics
2477	Politics
2487	Politics
2491	Politics
2492	Politics
2493	Politics
2496	Politics
2499	Politics
2507	Politics
2514	Politics
2526	Politics
2530	Politics
2537	Politics
2538	Politics
2539	Politics
2541	Politics
2542	Technology
2543	Technology
2545	Technology
2551	Film
2554	Film
2556	Business
2558	Business

2559	Politics
2560	Business
2561	Film
2566	Film
2568	Film
2573	Film
2574	Film
2575	Film
2580	Film
2583	Film
2591	Business
2597	Film
2599	Politics
2605	Politics
2607	Politics
2614	Film
2620	Politics
2622	Politics
2629	Politics
2641	Politics
2642	Politics
2643	Politics
2644	Politics
2649	Politics
2653	Business
2661	Business
2662	Politics
2667	Business
2670	Politics
2678	Politics
2681	Politics
2684	Politics
2691	Politics
2702	Business
2704	Technology
2706	Technology
2722	Politics
2732	Technology
2735	Business
2736	Politics

2741	Business
2745	Politics
2747	Business
2748	Politics
2751	Business
2753	Politics
2760	Business
2763	Politics
2766	Politics
2768	Politics
2771	Politics
2774	Film
2777	Politics
2778	Politics
2781	Politics
2782	Business
2783	Politics
2788	Politics
2789	Politics
2794	Politics
2798	Politics
2804	Politics
2807	Business
2808	Politics
2812	Politics
2818	Politics
2820	Politics
2833	Politics
2834	Politics
2837	Politics
2840	Politics
2841	Politics
2842	Politics
2849	Business
2855	Business
2861	Politics
2870	Politics
2873	Politics
2878	Business
2885	Business

2892	Politics
2896	Politics
2905	Politics
2908	Business
2916	Politics
2917	Politics
2926	Business
2927	Politics
2932	Politics
2940	Politics
2944	Politics
2946	Politics
2947	Politics
2948	Business
2952	Politics
2957	Politics
2969	Politics
2970	Business
2979	Technology
2994	Business
2998	Technology
2999	Technology
3002	Technology
3003	Technology
3009	Technology
3013	Technology
3015	Politics
3016	Politics
3021	Politics
3023	Technology
3029	Technology
3030	Politics
3032	Politics
3033	Business
3039	Business
3048	Business
3054	Film
3055	Business
3065	Business
3066	Business

3072	Business
3075	Business
3077	Business
3078	Business
3079	Business
3081	Business
3094	Business
3095	Politics
3105	Business
3108	Politics
3110	Business
3117	Business
3121	Business
3126	Business
3128	Business
3129	Business
3130	Business
3131	Business
3132	Business
3134	Business
3135	Business
3144	Business
3173	Business
3175	Business
3178	Business
3185	Business
3187	Business
3211	Business
3212	Technology
3221	Business
3223	Business
3225	Business
3229	Business
3232	Business
3236	Business
3248	Business
3258	Business
3262	Business
3266	Business
3268	Business

3277	Business
3279	Business
3284	Business
3286	Business
3289	Business
3291	Business
3293	Business
3294	Business
3298	Film
3299	Film
3309	Politics
3321	Politics
3323	Politics
3329	Business
3334	Business
3335	Business
3345	Politics
3347	Business
3355	Business
3359	Business
3367	Business
3369	Business
3375	Business
3379	Business
3382	Business
3390	Business
3391	Technology
3396	Business
3398	Business
3401	Business
3405	Politics
3414	Politics
3416	Politics
3419	Politics
3420	Business
3429	Business
3432	Business
3440	Business
3441	Business
3442	Business

3448	Politics
3457	Business
3462	Business
3463	Business
3464	Business
3472	Business
3474	Business
3477	Business
3478	Business
3485	Business
3488	Business
3501	Business
3506	Business
3521	Politics
3528	Business
3534	Politics
3543	Politics
3554	Politics
3555	Politics
3559	Business
3561	Politics
3563	Politics
3575	Politics
3577	Politics
3580	Technology
3582	Politics
3594	Technology
3607	Technology
3614	Film
3630	Technology
3636	Technology
3638	Technology
3643	Technology
3644	Technology
3647	Technology
3649	Technology
3655	Film
3656	Film
3667	Technology
3668	Technology

3672	Technology
3673	Technology
3675	Technology
3676	Technology
3677	Technology
3686	Technology
3688	Business
3690	Technology
3693	Technology
3697	Technology
3699	Technology
3705	Technology
3711	Technology
3713	Technology
3719	Technology
3721	Technology
3726	Technology
3727	Technology
3733	Business
3738	Business
3740	Politics
3747	Business
3753	Business
3762	Politics
3768	Business
3770	Technology
3772	Technology
3774	Technology
3775	Technology
3787	Technology
3792	Technology
3794	Technology
3795	Technology
3819	Football
3820	Football
3837	Football
3840	Football
3846	Football
3851	Business
3855	Football

3857	Business
3863	Business
3869	Business
3874	Football
3880	Politics
3881	Politics
3885	Football
3889	Football
3893	Football
3896	Football
3898	Football
3903	Football
3904	Football
3905	Football
3910	Football
3914	Politics
3919	Football
3924	Politics
3927	Football
3929	Politics
3949	Football
3951	Football
3952	Business
3959	Football
3963	Football
3974	Football
3979	Football
3983	Football
3985	Football
3990	Football
3991	Football
3992	Football
3994	Football
3997	Football
3999	Football
4006	Technology
4015	Football
4021	Football
4022	Football
4031	Football

4040	Football
4045	Politics
4047	Football
4060	Football
4066	Football
4067	Football
4072	Football
4078	Football
4086	Football
4092	Business
4098	Business
4106	Business
4107	Business
4109	Business
4110	Technology
4114	Business
4116	Business
4121	Technology
4122	Technology
4124	Technology
4125	Politics
4126	Technology
4128	Technology
4130	Technology
4132	Technology
4144	Technology
4155	Technology
4163	Technology
4165	Technology
4168	Technology
4176	Technology
4184	Technology
4187	Technology
4195	Technology
4200	Technology
4213	Technology
4222	Technology
4223	Technology
4230	Technology
4235	Technology

4236	Technology
4242	Technology
4243	Technology
4258	Film
4261	Technology
4262	Technology
4263	Technology
4266	Technology
4268	Technology
4279	Technology
4297	Technology
4299	Technology
4300	Technology
4301	Technology
4309	Technology
4320	Technology
4327	Technology
4330	Technology
4338	Technology
4340	Technology
4341	Technology
4352	Technology
4353	Technology
4356	Technology
4357	Technology
4361	Technology
4374	Technology
4383	Technology
4387	Technology
4400	Technology
4402	Technology
4414	Technology
4416	Football
4421	Football
4423	Football
4426	Football
4427	Football
4428	Football
4432	Football
4435	Football

4437	Football
4441	Football
4451	Football
4453	Football
4456	Football
4457	Football
4461	Football
4462	Football
4465	Football
4466	Football
4469	Football
4482	Football
4496	Football
4497	Football
4502	Football
4503	Football
4510	Football
4513	Football
4518	Football
4523	Football
4528	Football
4533	Football
4540	Football
4549	Football
4554	Football
4555	Football
4566	Football
4569	Football
4588	Football
4593	Football
4600	Football
4601	Football
4603	Football
4608	Football
4611	Football
4613	Football
4616	Football
4617	Football
4619	Football
4621	Football

4632	Football
4635	Football
4640	Football
4648	Football
4659	Football
4662	Football
4664	Football
4669	Football
4674	Football
4677	Football
4682	Football
4683	Football
4685	Football
4687	Football
4689	Football
4699	Football
4713	Football
4717	Football
4722	Football
4730	Football
4732	Football
4743	Football
4744	Football
4745	Football
4746	Football
4748	Football
4751	Football
4756	Football
4757	Football
4758	Football
4760	Football
4761	Football
4771	Football
4773	Football
4776	Football
4781	Football
4784	Football
4798	Football
4800	Football
4805	Football

4809	Football
4817	Football
4824	Football
4825	Football
4828	Football
4830	Football
4833	Football
4837	Football
4838	Football
4847	Football
4848	Football
4854	Football
4855	Football
4861	Football
4864	Football
4865	Football
4888	Football
4890	Football
4898	Football
4900	Football
4902	Football
4910	Football
4912	Football
4914	Football
4916	Football
4920	Football
4923	Football
4926	Football
4927	Football
4928	Football
4933	Football
4940	Football
4943	Football
4963	Football
4966	Football
4970	Football
4973	Football
4974	Football
4976	Football
4977	Football

4979	Football
4985	Football
4991	Football
4997	Football
5008	Football
5013	Football
5019	Football
5034	Football
5037	Football
5041	Football
5042	Football
5046	Football
5051	Football
5054	Football
5055	Football
5063	Football
5072	Football
5075	Football
5077	Football
5079	Football
5092	Football
5101	Football
5106	Football
5107	Football
5128	Football
5129	Football
5130	Football
5133	Football
5138	Football
5164	Football
5168	Football
5172	Business
5177	Football
5188	Football
5194	Technology
5195	Technology
5200	Football
5202	Technology
5210	Technology
5212	Technology

5214	Technology
5222	Business
5232	Technology
5237	Technology
5240	Technology
5247	Technology
5251	Film
5255	Technology
5256	Technology
5261	Business
5278	Technology
5281	Technology
5283	Film
5284	Business
5288	Technology
5290	Technology
5293	Football
5295	Technology
5299	Technology
5307	Technology
5311	Technology
5313	Football
5316	Business
5325	Technology
5338	Technology
5346	Technology
5348	Business
5350	Technology
5352	Technology
5360	Technology
5365	Technology
5368	Technology
5382	Technology
5390	Technology
5401	Technology
5407	Technology
5412	Technology
5419	Technology
5420	Business
5424	Technology

5427	Technology
5432	Film
5435	Technology
5439	Technology
5440	Technology
5441	Technology
5448	Technology
5451	Technology
5452	Technology
5455	Technology
5458	Business
5462	Technology
5463	Film
5464	Technology
5466	Technology
5471	Technology
5478	Film
5479	Technology
5482	Technology
5483	Technology
5491	Film
5492	Film
5499	Film
5500	Film
5503	Film
5513	Film
5514	Film
5518	Film
5522	Film
5527	Film
5528	Film
5530	Film
5536	Technology
5541	Film
5550	Film
5553	Film
5554	Film
5556	Film
5561	Film
5564	Film

5565	Film
5570	Film
5574	Film
5581	Film
5582	Film
5586	Film
5592	Film
5599	Film
5610	Film
5615	Film
5622	Film
5623	Film
5628	Film
5637	Film
5639	Business
5641	Film
5648	Film
5649	Film
5660	Film
5662	Film
5664	Film
5666	Film
5667	Film
5679	Film
5683	Film
5689	Film
5692	Technology
5693	Film
5694	Film
5700	Film
5701	Film
5704	Film
5707	Film
5709	Film
5711	Film
5714	Film
5717	Film
5719	Film
5728	Film
5736	Film

5738	Film
5742	Film
5744	Film
5751	Film
5752	Film
5757	Film
5760	Film
5761	Film
5763	Film
5775	Film
5781	Film
5784	Film
5787	Film
5796	Film
5804	Film
5805	Film
5816	Film
5821	Film
5823	Film
5827	Film
5829	Film
5833	Film
5837	Film
5838	Film
5842	Film
5849	Film
5851	Film
5865	Film
5873	Film
5886	Film
5895	Film
5900	Film
5906	Film
5910	Film
5913	Business
5914	Business
5918	Film
5930	Politics
5937	Politics
5942	Business

5943	Business
5947	Business
5955	Business
5960	Business
5962	Business
5964	Politics
5968	Business
5971	Business
5977	Business
5981	Politics
5989	Business
5990	Business
5994	Business
5998	Business
6002	Business
6007	Business
6018	Business
6019	Business
6020	Business
6026	Business
6047	Business
6056	Business
6058	Technology
6060	Business
6062	Business
6067	Politics
6069	Technology
6071	Business
6073	Technology
6074	Technology
6081	Business
6083	Technology
6085	Business
6086	Politics
6087	Business
6093	Business
6101	Politics
6108	Business
6110	Football
6111	Football

6115	Football
6120	Business
6121	Technology
6125	Technology
6126	Football
6128	Politics
6130	Business
6131	Technology
6139	Technology
6140	Business
6143	Business
6157	Technology
6170	Politics
6176	Politics
6179	Technology
6180	Technology
6186	Technology
6187	Technology
6188	Technology
6196	Technology
6217	Football
6227	Football
6228	Football
6229	Football
6234	Football
6236	Football
6244	Technology
6249	Technology
6257	Football
6259	Football
6264	Football
6268	Football
6272	Film
6280	Football
6288	Football
6289	Football
6301	Football
6304	Football
6305	Football
6315	Football

6316	Football
6320	Football
6321	Football
6333	Football
6334	Football
6341	Football
6342	Technology
6347	Business
6350	Business
6352	Football
6356	Technology
6357	Technology
6358	Film
6381	Film
6394	Film
6416	Film
6418	Business
6422	Business
6424	Business
6425	Business
6427	Technology
6431	Politics
6432	Politics
6442	Politics
6443	Politics
6445	Politics
6456	Business
6459	Football
6460	Business
6473	Technology
6477	Film
6479	Technology
6482	Film
6499	Film
6501	Technology
6515	Technology
6517	Film
6518	Technology
6521	Film
6522	Technology

6530	Film
6534	Film
6536	Film
6538	Technology
6542	Film
6546	Technology
6548	Film
6558	Film
6568	Technology
6575	Business
6578	Business
6581	Business
6590	Business
6609	Politics
6611	Politics
6612	Politics
6624	Business
6629	Business
6630	Business
6632	Technology
6634	Politics
6637	Business
6642	Business
6643	Business
6644	Business
6657	Business
6660	Business
6664	Business
6679	Politics
6680	Business
6685	Business
6703	Business
6705	Business
6706	Politics
6707	Politics
6708	Business
6709	Business
6710	Business
6711	Business
6712	Business

6716	Business
6723	Business
6725	Business
6738	Business
6742	Business
6755	Business
6757	Business
6761	Business
6762	Business
6763	Business
6764	Business
6769	Business
6772	Business
6773	Business
6777	Business
6778	Business
6788	Politics
6792	Business
6794	Business
6800	Business
6806	Politics
6820	Business
6824	Business
6826	Business
6829	Business
6832	Business
6837	Business
6840	Business
6841	Business
6847	Business
6857	Business
6868	Technology
6882	Technology
6893	Business
6895	Business
6897	Business
6898	Politics
6907	Business
6911	Business
6917	Film

6939	Business
6943	Technology
6952	Technology
6953	Politics
6954	Politics
6959	Politics
6965	Technology
6974	Technology
6982	Technology
6987	Technology
7001	Football
7003	Business
7006	Football
7007	Football
7012	Business
7013	Business
7020	Football
7024	Film
7030	Business
7031	Business
7035	Business
7038	Business
7044	Technology
7045	Technology
7049	Football
7051	Business
7057	Technology
7058	Business
7059	Business
7061	Technology
7068	Business
7071	Business
7090	Technology
7101	Technology
7105	Technology
7108	Technology
7109	Politics
7110	Politics
7114	Technology
7121	Technology

7123	Technology
7124	Politics
7126	Technology
7132	Technology
7133	Politics
7136	Technology
7141	Politics
7142	Politics
7146	Politics
7147	Politics
7151	Politics
7153	Politics
7154	Business
7157	Politics
7163	Film
7166	Film
7167	Film
7174	Film
7183	Business
7186	Business
7194	Technology
7205	Business
7214	Business
7216	Football
7221	Football
7225	Film
7234	Film
7240	Film
7248	Film
7251	Politics
7267	Football
7270	Film
7283	Football
7288	Football
7291	Football
7318	Business
7324	Business
7325	Business
7327	Business
7328	Technology

7329	Business
7331	Business
7345	Politics
7362	Politics
7367	Film
7369	Film
7370	Politics
7371	Politics
7376	Technology
7377	Technology
7390	Business
7396	Business
7401	Politics
7411	Business
7425	Politics
7426	Business
7445	Politics
7454	Politics
7455	Politics
7457	Politics
7459	Politics
7461	Politics
7469	Football
7471	Football
7478	Politics
7488	Politics
7490	Politics
7491	Business
7496	Politics
7514	Politics
7515	Politics
7521	Business
7525	Politics
7532	Politics
7548	Business
7552	Business
7553	Politics
7556	Politics
7566	Politics
7570	Politics

7576	Business
7577	Politics
7590	Business
7601	Politics
7606	Politics
7609	Politics
7612	Politics
7619	Politics
7633	Politics
7641	Politics
7643	Politics
7645	Politics
7661	Politics
7666	Politics
7668	Politics
7671	Politics
7676	Politics
7677	Politics
7684	Politics
7693	Politics
7694	Politics
7707	Politics
7710	Politics
7712	Politics
7716	Politics
7717	Politics
7721	Politics
7724	Politics
7735	Politics
7737	Football
7745	Politics
7757	Politics
7760	Football
7762	Business
7764	Politics
7765	Politics
7770	Business
7772	Politics
7773	Politics
7774	Politics

7777	Politics
7783	Business
7789	Football
7790	Politics
7800	Politics
7804	Business
7813	Politics
7814	Politics
7816	Politics
7822	Politics
7823	Business
7833	Politics
7836	Politics
7840	Politics
7844	Politics
7846	Business
7848	Politics
7855	Technology
7857	Politics
7871	Business
7877	Business
7881	Politics
7884	Business
7896	Politics
7901	Politics
7920	Politics
7926	Business
7929	Business
7933	Business
7937	Business
7941	Business
7946	Business
7948	Business
7954	Business
7960	Business
7964	Business
7968	Politics
7973	Business
7975	Business
7978	Business

7982	Business
7993	Business
7994	Business
7998	Business
8001	Business
8011	Business
8016	Politics
8022	Politics
8036	Politics
8040	Business
8044	Business
8045	Business
8046	Politics
8054	Business
8064	Business
8067	Business
8083	Business
8086	Politics
8090	Business
8095	Business
8100	Politics
8103	Politics
8106	Politics
8107	Business
8112	Politics
8113	Politics
8120	Politics
8121	Business
8125	Politics
8134	Business
8140	Business
8141	Business
8142	Business
8144	Politics
8145	Politics
8149	Business
8152	Business
8154	Business
8156	Business
8158	Business

8160	Business
8165	Business
8171	Business
8175	Business
8176	Business
8178	Business
8179	Politics
8181	Business
8183	Business
8185	Politics
8188	Business
8194	Business
8199	Politics
8205	Business
8208	Politics
8209	Business
8211	Business
8212	Business
8216	Business
8224	Business
8228	Business
8231	Business
8234	Business
8236	Business
8241	Business
8246	Business
8250	Business
8262	Business
8268	Business
8272	Business
8276	Business
8277	Business
8284	Politics
8287	Politics
8295	Business
8301	Business
8332	Business
8338	Business
8346	Business
8347	Business

8348	Business
8359	Business
8360	Film
8361	Film
8367	Business
8368	Business
8372	Business
8377	Business
8380	Business
8385	Football
8387	Technology
8401	Business
8405	Film
8408	Business
8415	Politics
8417	Business
8421	Politics
8429	Politics
8438	Politics
8447	Politics
8448	Politics
8454	Business
8464	Business
8467	Politics
8468	Business
8479	Business
8482	Business
8483	Business
8484	Technology
8486	Business
8511	Business
8515	Business
8537	Politics
8559	Politics
8563	Business
8575	Politics
8577	Politics
8580	Business
8581	Business
8582	Politics

8583	Politics
8590	Politics
8595	Politics
8597	Business
8599	Politics
8611	Politics
8614	Politics
8615	Politics
8620	Politics
8622	Politics
8624	Business
8625	Business
8626	Politics
8628	Politics
8639	Business
8650	Business
8651	Business
8656	Business
8659	Business
8660	Business
8669	Business
8677	Business
8697	Politics
8705	Business
8709	Business
8711	Politics
8713	Politics
8715	Business
8724	Politics
8728	Business
8730	Business
8731	Business
8733	Business
8739	Business
8743	Politics
8751	Business
8752	Business
8753	Politics
8755	Politics
8759	Business

8760	Business
8766	Business
8768	Business
8775	Business
8779	Politics
8798	Politics
8800	Politics
8801	Business
8803	Business
8810	Business
8812	Business
8813	Business
8814	Politics
8818	Politics
8821	Business
8825	Business
8826	Business
8828	Politics
8831	Politics
8832	Politics
8835	Business
8840	Politics
8857	Politics
8860	Politics
8865	Politics
8872	Football
8877	Technology
8885	Football
8888	Technology
8891	Politics
8903	Technology
8904	Technology
8906	Technology
8918	Technology
8931	Technology
8932	Technology
8938	Film
8941	Football
8946	Technology
8948	Football

8950	Technology
8951	Football
8952	Football
8956	Technology
8960	Football
8964	Technology
8972	Football
8978	Film
8982	Technology
8983	Technology
8990	Football
8991	Politics
9001	Business
9010	Business
9014	Technology
9019	Politics
9032	Politics
9047	Politics
9058	Politics
9060	Politics
9066	Business
9083	Business
9086	Business
9097	Politics
9099	Business
9100	Politics
9110	Business
9114	Politics
9134	Business
9140	Business
9145	Politics
9146	Politics
9152	Politics
9159	Politics
9173	Film
9175	Film
9180	Business
9181	Business
9186	Film
9188	Film

9200	Film
9202	Business
9208	Business
9213	Film
9216	Film
9220	Film
9224	Film
9232	Film
9235	Film
9238	Film
9244	Film
9254	Technology
9257	Film
9264	Film
9274	Politics
9283	Politics
9284	Politics
9287	Politics
9288	Film
9292	Film
9293	Politics
9294	Politics
9296	Politics
9297	Film
9307	Politics
9308	Politics
9310	Politics
9317	Film
9318	Film
9319	Film
9325	Politics
9333	Film
9334	Politics
9336	Politics
9341	Politics
9345	Film
9347	Politics
9359	Film
9360	Film
9361	Film

9364	Film
9368	Film
9380	Film
9385	Film
9389	Film
9391	Film
9395	Film
9397	Film
9420	Film
9424	Film
9426	Politics
9432	Politics
9433	Film
9437	Film
9438	Politics
9445	Film
9491	Politics
9494	Film
9495	Film
9519	Film
9523	Film
9525	Film
9528	Technology
9531	Film
9543	Film
9545	Film
9546	Film
9550	Politics
9567	Film
9572	Film
9580	Film
9582	Film
9583	Film
9584	Politics
9587	Film
9593	Film
9594	Film
9595	Film
9596	Film
9609	Film

9611	Film
9612	Film
9616	Film
9621	Film
9622	Film
9624	Film
9629	Film
9634	Film
9639	Film
9647	Film
9648	Film
9649	Film
9652	Film
9656	Film
9658	Film
9662	Film
9677	Film
9679	Film
9682	Film
9685	Film
9697	Film
9703	Film
9713	Film
9725	Politics
9732	Politics
9759	Technology
9774	Film
9775	Film
9779	Film
9781	Film
9782	Film
9795	Football
9804	Business
9806	Politics
9817	Business
9820	Football
9823	Politics
9830	Business
9843	Film
9846	Film

9850	Technology
9855	Business
9859	Politics
9869	Business
9887	Business
9890	Politics
9892	Business
9899	Business
9902	Politics
9903	Politics
9905	Politics
9915	Film
9918	Film
9921	Film
9931	Technology
9940	Business
9942	Film
9945	Politics
9949	Politics
9950	Politics
9957	Politics
9958	Politics
9961	Politics
9987	Politics
9991	Politics
9994	Politics
9997	Politics
10005	Business
10008	Politics
10009	Politics
10010	Politics
10012	Politics
10016	Politics
10017	Business
10041	Film
10042	Film
10051	Film
10061	Film
10065	Film
10068	Film

10078	Film
10081	Film
10094	Film
10095	Film
10100	Business
10109	Football
10113	Film
10117	Film
10118	Film
10123	Film
10129	Film
10133	Film
10147	Film
10149	Football
10152	Film
10157	Film
10161	Film
10169	Film
10175	Politics
10177	Film
10186	Business
10191	Football
10203	Football
10207	Football
10208	Football
10214	Business
10215	Football
10216	Politics
10221	Business
10232	Football
10236	Politics
10237	Business
10244	Football
10262	Football
10263	Football
10268	Football
10269	Football
10270	Football
10272	Film
10279	Football

10286	Football
10288	Football
10290	Football
10291	Football
10293	Football
10294	Football
10296	Football
10306	Football
10309	Football
10310	Football
10311	Football
10315	Football
10316	Football
10327	Football
10329	Football
10334	Football
10336	Football
10337	Football
10344	Business
10355	Football
10366	Football
10378	Football
10383	Football
10389	Football
10391	Football
10392	Football
10399	Football
10407	Football
10412	Football
10422	Football
10427	Football
10429	Football
10432	Football
10440	Football
10443	Football
10444	Football
10448	Football
10453	Football
10460	Football
10462	Football

10463 Football
10464 Football
10466 Football
10471 Football
10474 Football
10480 Football
10481 Football
10483 Football
10486 Football
10494 Football
10496 Football
10497 Football
10510 Football
10516 Football
10520 Football
10524 Football
10535 Football
10538 Football
10540 Football
10545 Technology
10546 Football
10547 Football
10552 Football
10575 Football
10579 Football
10587 Football
10588 Football
10593 Football
10597 Football
10598 Football
10617 Football
10628 Football
10629 Football
10633 Film
10635 Football
10638 Football
10644 Football
10645 Football
10657 Football
10669 Politics

10670	Football
10671	Football
10674	Business
10694	Business
10697	Football
10698	Business
10701	Business
10702	Football
10705	Football
10706	Business
10709	Football
10713	Football
10715	Business
10720	Football
10732	Football
10738	Football
10740	Football
10746	Football
10751	Business
10758	Football
10759	Football
10762	Football
10771	Business
10773	Football
10781	Politics
10782	Football
10783	Football
10797	Business
10799	Politics
10803	Business
10806	Business
10813	Film
10819	Business
10825	Business
10831	Business
10834	Football
10837	Politics
10844	Football
10854	Politics
10856	Football

10858 Technology
10861 Politics
10867 Business
10871 Business
10885 Business
10892 Football
10899 Politics
10905 Politics
10907 Film
10918 Film
10920 Film
10926 Politics
10929 Politics
10931 Film
10948 Film
10949 Technology
10952 Film
10954 Film
10956 Film
10962 Business
10970 Politics
10972 Politics
10981 Film
10987 Film
10999 Football
11003 Football
11014 Football
11019 Football
11026 Football
11033 Football
11036 Business
11037 Football
11039 Football
11041 Politics
11042 Football
11045 Business
11051 Politics
11052 Politics
11071 Politics
11073 Politics

11097 Business
11098 Business
11099 Business
11115 Business
11132 Politics
11141 Politics
11147 Politics
11152 Business
11154 Politics
11169 Politics
11170 Politics
11180 Business
11181 Politics
11185 Politics
11188 Business
11190 Politics
11192 Politics
11203 Business
11204 Business
11210 Politics
11216 Business
11218 Business
11221 Politics
11230 Politics
11231 Politics
11237 Politics
11252 Business
11257 Politics
11259 Politics
11266 Business
11270 Politics
11271 Politics
11273 Politics
11278 Politics
11290 Politics
11297 Business
11302 Politics
11307 Politics
11308 Politics
11315 Politics

11316 Politics
11319 Politics
11320 Politics
11331 Politics
11335 Film
11336 Politics
11344 Politics
11354 Politics
11356 Business
11359 Politics
11386 Politics
11387 Politics
11395 Politics
11397 Business
11403 Business
11405 Politics
11419 Politics
11439 Business
11447 Politics
11448 Politics
11451 Business
11458 Politics
11464 Politics
11467 Politics
11470 Politics
11483 Politics
11489 Politics
11490 Politics
11494 Politics
11502 Politics
11506 Politics
11509 Football
11510 Politics
11513 Football
11515 Football
11520 Football
11526 Football
11530 Football
11532 Football
11535 Football

11554 Football
11560 Football
11565 Football
11569 Football
11571 Football
11575 Football
11577 Football
11582 Football
11585 Football
11587 Football
11589 Football
11594 Football
11595 Football
11602 Football
11607 Football
11610 Football
11613 Football
11619 Football
11624 Business
11631 Football
11634 Football
11643 Football
11646 Football
11649 Football
11650 Football
11653 Football
11658 Football
11665 Football
11670 Football
11674 Football
11678 Football
11683 Football
11695 Football
11702 Football
11708 Football
11710 Football
11715 Football
11723 Football
11727 Football
11740 Football

11756 Football
11758 Football
11759 Football
11761 Football
11765 Football
11769 Football
11771 Football
11779 Football
11785 Football
11795 Football
11796 Football
11802 Football
11804 Football
11814 Football
11821 Football
11824 Football
11833 Football
11835 Football
11845 Football
11852 Football
11859 Football
11867 Football
11869 Football
11876 Football
11878 Football
11890 Football
11892 Football
11900 Football
11904 Football
11908 Football
11910 Football
11918 Football
11923 Football
11930 Football
11933 Football
11934 Football
11938 Football
11940 Football
11941 Football
11943 Football

11944 Football
11946 Football
11959 Football
11980 Football
11986 Football
11989 Football
11998 Football
12002 Football
12003 Football
12005 Football
12017 Football
12030 Football
12035 Football
12036 Football
12038 Football
12041 Football
12042 Film
12043 Football
12048 Football
12060 Film
12073 Film
12078 Film
12082 Film
12083 Football
12084 Football
12085 Football
12100 Film
12104 Film
12118 Film
12122 Technology
12127 Film
12134 Football
12135 Football
12138 Football
12140 Film
12157 Film
12164 Film
12166 Film
12167 Film
12172 Film

12190	Film
12197	Football
12202	Football
12204	Technology
12206	Technology
12207	Technology
12212	Football
12213	Football
12215	Film
12218	Technology
12223	Technology
12226	Football
12228	Football
12229	Football
12240	Technology
12241	Football
12256	Football
12266	Football
12270	Football
12277	Football
12279	Football
12283	Football
12284	Technology
12286	Technology
12295	Football
12307	Football
12320	Football
12321	Football
12322	Football
12326	Football
12327	Football
12329	Football
12331	Football
12337	Football
12338	Football
12342	Film
12343	Football
12350	Football
12374	Football
12376	Football

12378	Football
12384	Football
12387	Football
12388	Football
12389	Football
12393	Football
12399	Football
12409	Football
12417	Football
12420	Football
12427	Football
12439	Football
12440	Football
12448	Football
12458	Football
12466	Football
12470	Football
12476	Football
12491	Football
12495	Football
12502	Football
12507	Football
12510	Football
12519	Football
12521	Football
12525	Football
12526	Football
12530	Football
12534	Football
12535	Football
12537	Politics
12538	Football
12539	Football
12543	Football
12545	Football
12549	Football
12550	Football
12556	Politics
12562	Football
12568	Politics

12573 Football
12574 Politics
12581 Politics
12583 Politics
12585 Politics
12586 Football
12587 Football
12588 Politics
12589 Football
12591 Football
12597 Politics
12599 Politics
12600 Politics
12602 Football
12603 Politics
12625 Politics
12630 Film
12665 Business
12666 Business
12674 Film
12688 Business
12697 Film
12698 Film
12705 Film
12715 Politics
12718 Film
12724 Politics
12734 Film
12736 Film
12740 Film
12743 Film
12753 Film
12758 Film
12764 Film
12776 Film
12777 Film
12779 Film
12785 Film
12787 Film
12791 Film

12793	Film
12796	Film
12800	Film
12810	Film
12814	Film
12817	Film
12818	Film
12820	Film
12824	Film
12829	Film
12835	Technology
12851	Film
12855	Film
12856	Film
12857	Film
12865	Film
12884	Film
12885	Film
12893	Film
12897	Film
12904	Film
12909	Film
12913	Business
12916	Film
12920	Film
12922	Film
12934	Film
12940	Film
12947	Film
12950	Film
12952	Film
12960	Film
12962	Film
12974	Film
12976	Film
12979	Film
13009	Film
13011	Film
13020	Film
13028	Film

13039 Film
13041 Film
13042 Film
13044 Film
13045 Film
13048 Film
13052 Business
13055 Business
13056 Business
13057 Film
13058 Film
13077 Film
13086 Film
13091 Film
13095 Film
13096 Film
13109 Technology
13111 Technology
13112 Technology
13120 Film
13122 Film
13132 Business
13133 Technology
13135 Film
13139 Technology
13142 Business
13144 Film
13148 Technology
13157 Business
13158 Business
13162 Football
13164 Technology
13166 Football
13171 Football
13173 Football
13174 Politics
13186 Film
13187 Politics
13188 Football
13190 Politics

13191 Politics
13204 Politics
13207 Football
13210 Football
13218 Football
13225 Football
13231 Football
13236 Football
13247 Football
13250 Football
13254 Football
13256 Football
13257 Football
13262 Football
13263 Football
13268 Football
13272 Football
13273 Football
13283 Football
13284 Football
13297 Football
13299 Football
13304 Football
13307 Football
13309 Football
13313 Football
13316 Football
13317 Football
13339 Football
13349 Football
13352 Football
13357 Football
13358 Football
13364 Football
13371 Football
13381 Football
13394 Football
13399 Politics
13401 Football
13402 Football

13404 Football
13407 Politics
13427 Football
13430 Football
13445 Football
13453 Football
13455 Football
13463 Football
13480 Film
13485 Football
13487 Football
13494 Football
13500 Film
13505 Football
13508 Football
13509 Football
13522 Football
13526 Football
13527 Football
13538 Football
13542 Football
13543 Football
13551 Football
13567 Football
13569 Football
13574 Football
13575 Football
13578 Football
13583 Football
13588 Football
13590 Football
13609 Football
13610 Football
13613 Football
13616 Football
13621 Football
13622 Football
13623 Football
13625 Football
13627 Football

13631 Football
13633 Football
13634 Football
13643 Football
13647 Football
13648 Football
13657 Football
13661 Football
13668 Football
13675 Football
13678 Football
13679 Football
13684 Football
13687 Football
13691 Football
13697 Football
13701 Football
13709 Football
13711 Football
13715 Football
13717 Football
13721 Football
13729 Film
13733 Football
13738 Football
13753 Football
13760 Football
13768 Football
13769 Football
13770 Football
13774 Football
13775 Football
13776 Football
13777 Football
13782 Football
13785 Football
13801 Football
13803 Football
13807 Football
13809 Technology

13823 Technology
13826 Technology
13827 Technology
13836 Technology
13847 Technology
13849 Technology
13851 Technology
13852 Football
13857 Technology
13871 Technology
13875 Business
13877 Technology
13879 Business
13890 Technology
13897 Business
13905 Business
13918 Technology
13925 Business
13938 Football
13942 Football
13946 Football
13954 Football
13957 Technology
13958 Football
13961 Politics
13970 Football
13975 Football
13977 Football
13978 Football
13985 Business
13997 Football
13998 Football
14004 Football
14010 Football
14015 Football
14016 Football
14020 Football
14022 Football
14025 Football
14028 Football

14030 Football
14038 Football
14045 Football
14054 Business
14056 Business
14059 Technology
14066 Politics
14068 Football
14071 Business
14076 Politics
14077 Technology
14082 Business
14095 Politics
14105 Politics
14108 Business
14113 Business
14116 Politics
14130 Politics
14131 Business
14141 Technology
14145 Technology
14148 Technology
14152 Technology
14153 Football
14163 Film
14164 Film
14179 Business
14182 Football
14184 Business
14196 Film
14199 Film
14200 Film
14209 Politics
14211 Politics
14212 Politics
14217 Politics
14222 Politics
14223 Politics
14228 Politics
14241 Politics

14243 Politics
14246 Politics
14253 Politics
14260 Politics
14261 Politics
14268 Politics
14270 Business
14271 Politics
14279 Politics
14280 Business
14282 Politics
14288 Politics
14289 Politics
14290 Film
14300 Business
14304 Business
14309 Business
14313 Business
14322 Business
14323 Business
14324 Business
14327 Politics
14328 Politics
14334 Business
14338 Business
14339 Business
14345 Business
14350 Politics
14352 Politics
14354 Politics
14358 Politics
14362 Technology
14366 Football
14370 Football
14373 Football
14380 Business
14386 Business
14387 Business
14388 Business
14391 Technology

14395 Football
14398 Football
14411 Football
14419 Football
14422 Football
14424 Football
14425 Football
14430 Football
14431 Football
14432 Football
14434 Football
14435 Football
14446 Football
14448 Football
14450 Football
14453 Football
14454 Football
14459 Football
14460 Football
14464 Technology
14477 Technology
14481 Technology
14483 Technology
14485 Business
14487 Technology
14488 Technology
14489 Film
14490 Politics
14494 Film
14495 Film
14496 Film
14504 Politics
14506 Business
14508 Business
14514 Politics
14516 Technology
14517 Technology
14525 Politics
14529 Business
14550 Politics

14551 Technology
14557 Technology
14572 Football
14579 Business
14581 Politics
14582 Politics
14588 Politics
14592 Politics
14594 Politics
14595 Business
14597 Politics
14602 Business
14604 Politics
14609 Politics
14616 Politics
14618 Politics
14621 Politics
14622 Politics
14624 Politics
14627 Football
14629 Politics
14631 Football
14632 Politics
14633 Politics
14643 Business
14645 Business
14649 Business
14652 Business
14653 Business
14659 Business
14660 Business
14676 Business
14677 Business
14679 Business
14683 Business
14684 Business
14685 Business
14689 Business
14690 Business
14694 Business

14695 Technology
14703 Business
14706 Business
14709 Politics
14712 Business
14713 Politics
14718 Politics
14720 Politics
14722 Business
14730 Football
14737 Politics
14740 Politics
14742 Politics
14743 Politics
14746 Business
14748 Politics
14750 Film
14755 Film
14756 Film
14762 Politics
14763 Film
14765 Film
14766 Film
14773 Film
14774 Film
14775 Film
14780 Film
14787 Politics
14795 Politics
14796 Technology
14799 Film
14801 Politics
14803 Politics
14807 Politics
14813 Film
14821 Film
14828 Football
14837 Football
14838 Football
14840 Football

14841 Football
14850 Football
14855 Business
14856 Technology
14859 Politics
14863 Business
14865 Politics
14868 Business
14879 Politics
14885 Business
14889 Politics
14891 Politics
14892 Business
14899 Business
14900 Football
14916 Business
14934 Football
14939 Football
14940 Football
14941 Football
14953 Football
14965 Football
14968 Football
14979 Football
14987 Football
14992 Football
15000 Football
15003 Football
15013 Football
15018 Football
15026 Football
15027 Football
15029 Football
15034 Film
15035 Football
15038 Football
15041 Football
15052 Technology
15054 Technology
15055 Film

15056 Football
15072 Football
15089 Football
15090 Football
15092 Football
15098 Football
15117 Politics
15119 Football
15125 Film
15126 Film
15129 Film
15139 Film
15143 Film
15154 Business
15170 Film
15174 Film
15178 Film
15179 Film
15183 Technology
15184 Technology
15186 Film
15200 Football
15201 Football
15221 Football
15224 Football
15225 Football
15228 Football
15234 Football
15236 Football
15239 Football
15249 Football
15251 Football
15256 Football
15260 Football
15264 Football
15267 Football
15269 Football
15272 Football
15273 Technology
15276 Business

15278	Technology
15284	Football
15290	Football
15298	Film
15299	Football
15302	Football
15304	Football
15305	Football
15315	Football
15321	Football
15322	Football
15325	Technology
15332	Football
15333	Football