1)
Top 100 most commonly used words in fake news:
trump: 74403
said: 31149
president: 26340
people: 26098
one: 23812
would: 23461
state: 22072
clinton: 18717
like: 18207
obama: 17920
time: 17885
donald: 17235
american: 16093
republican: 16061
say: 15528
also: 15243
year: 14843
new: 14198
news: 14198
u: 14172
image: 13937
even: 13692
hillary: 13678
white: 13146
right: 12698
get: 12230
know: 11947
make: 11534
via: 11355
woman: 11200
medium: 11142
campaign: 11069
house: 10774
country: 10770
america: 10703
could: 10230
first: 10041
want: 9818
think: 9765
going: 9750
many: 9719
way: 9394

election: 9297
day: 9217
told: 9103
government: 9079
thing: 8962
video: 8903
made: 8667
back: 8611
law: 8607
police: 8586
go: 8436
two: 8344
black: 8047
party: 8036
show: 8035
united: 7978
group: 7946
last: 7861
take: 7803
see: 7761
come: 7756
may: 7630
political: 7551
report: 7502
fact: 7379
well: 7215
national: 7204
need: 7151
former: 7124
vote: 7081
world: 6997
much: 6944
democrat: 6929
million: 6913
story: 6741
life: 6725
bill: 6571
public: 6539
official: 6427
support: 6410
according: 6332
man: 6329
attack: 6319
member: 6285

week: 6241
another: 6227
never: 6195
really: 6181
family: 6157
every: 6036
since: 6012
candidate: 5999
work: 5909
case: 5816
still: 5728
child: 5717
muslim: 5708
presidential: 5694
Top 100 most commonly used words in real news:
said: 99034
trump: 54269
state: 36232
would: 31524
reuters: 28412
president: 26928
republican: 22096
government: 19430
year: 18711
house: 16900
new: 16783
also: 15946
united: 15576
people: 15193
party: 14963
official: 14575
told: 14244
country: 13924
election: 13900
could: 13709
one: 13019
last: 12630
washington: 12418
two: 11619
group: 11103
campaign: 11074
former: 10601
leader: 10498
donald: 10447

week: 10419
security: 10374
court: 10336
percent: 9936
say: 9930
north: 9870
minister: 9541
white: 9500
clinton: 9499
tax: 9225
law: 9214
senate: 9204
obama: 9197
time: 9037
vote: 8976
month: 8754
china: 8562
first: 8547
national: 8533
statement: 8521
administration: 8375
since: 8332
tuesday: 8263
democratic: 8237
foreign: 8196
including: 8119
military: 8048
presidential: 8011
wednesday: 8008
democrat: 7943
right: 7849
russia: 7821
may: 7813
political: 7698
thursday: 7663
support: 7655
bill: 7579
million: 7562
policy: 7479
plan: 7382
friday: 7331
korea: 7257
day: 7172
monday: 7096

force: 7071
office: 6923
committee: 6874
member: 6834
american: 6822
deal: 6803
many: 6721
agency: 6527
senator: 6486
congress: 6484
federal: 6447
department: 6352
city: 6316
issue: 6271
company: 6212
made: 6201
make: 6151
according: 6142
part: 6141
comment: 6127
police: 6075
called: 6047
take: 6034
attack: 6010
russian: 6008
saying: 5985
news: 5973
Top 100 most commonly used words in all news:
said: 130183
trump: 128672
state: 58304
would: 54985
president: 53268
people: 41291
republican: 38157
one: 36831
year: 33554
also: 31189
new: 30981
reuters: 28799
government: 28509
clinton: 28216
donald: 27682
house: 27674

obama: 27117
time: 26922
say: 25458
country: 24694
could: 23939
united: 23554
told: 23347
election: 23197
party: 22999
american: 22915
like: 22833
white: 22646
campaign: 22143
official: 21002
right: 20547
last: 20491
news: 20171
two: 19963
group: 19049
first: 18588
washington: 17959
law: 17821
former: 17725
make: 17685
even: 17606
week: 16660
u: 16623
get: 16605
many: 16440
hillary: 16408
day: 16389
vote: 16057
security: 16048
court: 15856
national: 15737
want: 15582
medium: 15568
may: 15443
political: 15249
woman: 14873
democrat: 14872
made: 14868
leader: 14747
police: 14661

image: 14535
million: 14475
know: 14419
since: 14344
percent: 14172
bill: 14150
going: 14102
support: 14065
administration: 13979
think: 13910
take: 13837
way: 13780
back: 13768
presidential: 13705
statement: 13355
month: 13349
america: 13212
russia: 13144
member: 13119
democratic: 13012
tax: 12915
senate: 12726
policy: 12669
including: 12614
office: 12534
according: 12474
north: 12449
report: 12439
attack: 12329
need: 12015
department: 11991
public: 11895
via: 11864
go: 11804
federal: 11763
world: 11713
come: 11682
military: 11658
part: 11655
called: 11594

2)
Pure word-count analysis will not be able to show the difference between real or fake news. At least not by itself in this context

Too many words appear prominently in both sets of data, it would be extremely difficult for any human or machine to pick correctly based purely on that.
TFIDF improves upon this by comparing with other documents and could have some accuracy, but I believe the strongest feature set would be either bigrams or trigrams rather than the importance of singular words.

Problem 2)
Model 1 : SVM : Feature : Trigrams : Accuracy : 92.6% :
Model 2 : SVM : Feature : Bigrams : Precision : 98.43% : Accuracy : 97.75% : Recall : 96.86%
Model 3 : SVM : Feature : TFIDF : Precision : 94.46% : Accuracy : 91.44% : Recall : 91.41%

The confusion matrices revealed that both using TFIDF and Bigrams had the model guessing fake more often then real, with both more true positives and false positives for fake. However, both models vastly guessed true positives for both categories over false positives.

Problem 3)
Model 1 : SVM : Feature : Nouns and Adjectives Bigrrams : Precision: 86.57% : Accuracy: 78.60% : Recall: 65.83%
Model 3 : SVM : Feature : Nouns and Verbs Bigrams : Precision : 84.40% : Accuracy : 795.91% : Recall : 70.17%
Model 3 : SVM : Feature : Verbs and Adjectives Bigrams : Precision : 87.95% : Accuracy : 77.85% : Recall : 61.95%

Problem 4)
While bigrams seem to be a strong feature set for classifying fake news, to enhance the performance of the model I think I would utilize sentiment analysis instead of or in addition to bigrams, as a lot of fake news are intended to have emotional leanings. Utilizing a mixture of both bigrams and sentiment analysis could likely lead to a very robust classification model.

The easiest way to determine if GPT is viable for fake news detection is to use the API and pass the raw text article to GPT, with prompts designed to have GPT focus on detecting the validity of the article and reporting whether it believes the article to be real or fake. GPT has several strengths in this category due to its broad training set and contextual understanding, however is ultimately not viable - at least not without modification and re-training - due to the bias of any misinformation in its dataset (due to being trained on internet data) as well as not being directly trained on fact-checking news articles.