In [476]:
```python
import nltk
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from math import sqrt
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error
from math import sqrt
from sklearn.preprocessing import OneHotEncoder
from sklearn import preprocessing
from sklearn import linear_model
```

In [477]:
```python
import os
import re

import nltk

alice_file = 'alice.txt'
alice_raw = None

if not os.path.isfile(alice_file):
    from urllib import request
    url = 'http://www.gutenberg.org/cache/epub/19033/pg19033.txt'
    response = request.urlopen(url)
    alice_raw = response.read().decode('utf8')
    with open(alice_file, 'w', encoding='utf8') as f:
        f.write(alice_raw)
else:
    with open(alice_file, 'r', encoding='utf8') as f:
        alice_raw = f.read()
```

In [478]:
```python
# For reasons, lets remove the start and end bloat from the text
start = "I--DOWN THE RABBIT-HOLE"
end = "End of the Project Gutenberg"
start_index = alice_raw.find(start)
end_index = alice_raw.rfind(end)
alice = alice_raw[start_index:end_index]

# And replace more than one subsequent whitespace chars with one space
alice = re.sub(r'\s+', ' ', alice)
```

In [479]:
```python
# import WhitespaceTokenizer() method from nltk
from nltk.tokenize import WhitespaceTokenizer

# Create a reference variable for Class WhitespaceTokenizer
tk = WhitespaceTokenizer()

# Use tokenize method
tokens = tk.tokenize(alice)

train_data = tokens[:8691]
test_data = tokens[8691:]
```

In [553]:
```python
tokens_y = []
for word in tokens:
    y = ""
    for i in range (len(word)-1):
        y = y+'0'
    y = y+'1'
    tokens_y.append(y)
```

In [554]:
```python
x = ''.join(tokens)
y = ''.join(tokens_y)
```

In [574]:
```python
features = []
x_input = []
output = []
for i in range(len(x)):
    x_input.append(x[i])
    left = 0
    if i-5 > 0:
        left = i-5
    right = len(x)-1
    if i+5 < len(x):
        right = i+5
    xi = []
    for j in range(left, right+1):
        xi.append(x[j])
    features.append(xi)
    output.append(y[i])
```

In [588]:
```python
df = pd.DataFrame(features)
df.columns =['f1', 'f2', 'f3', 'f4', 'f5', 'f6', 'f7', 'f8', 'f9', 'f10
df['x'] = x_input
df['y'] = output
```

In [624]:
```
1  df
```

Out[624]:

	f1	f2	f3	f4	f5	f6	f7	f8	f9	f10	f11	x	y
0	l	-	-	D	O	W	None	None	None	None	None	l	0
1	l	-	-	D	O	W	N	None	None	None	None	-	0
2	l	-	-	D	O	W	N	T	None	None	None	-	0
3	l	-	-	D	O	W	N	T	H	None	None	D	0
4	l	-	-	D	O	W	N	T	H	E	None	O	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...
42509	u	s	t	r	a	t	i	o	n	]	None	t	0
42510	s	t	r	a	t	i	o	n	]	None	None	i	0
42511	t	r	a	t	i	o	n	]	None	None	None	o	0
42512	r	a	t	i	o	n	]	None	None	None	None	n	0
42513	a	t	i	o	n	]	None	None	None	None	None	]	1

42514 rows × 13 columns

Feaature Selction using one hot encoding

In [597]:
```
1  encoder = OneHotEncoder(handle_unknown='ignore', sparse=False)
2  onehot = encoder.fit_transform(df.loc[:,'f1':'f11'])
```

In [618]:
```
1  df_os = pd.DataFrame(onehot)
2  df_os.columns = [str(i) for i in range(onehot.shape[1])]
```

2/25/22, 11:19 PM

In [619]:
```
1  df.os
```

Out[619]:

	0	1	2	3	4	5	6	7	8	9	...	732	733	734	735	736	737	738	739
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
42509	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
42510	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
42511	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
42512	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
42513	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

42514 rows × 742 columns

In [620]:
```
1  df_combined = pd.concat([df, df_os], axis=1, join='inner')
2  df_combined
```

Out[620]:

	f1	f2	f3	f4	f5	f6	f7	f8	f9	f10	...	732	733	734	735	736	737	738	73
0	l	-	-	D	O	W	None	None	None	None	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.
1	l	-	-	D	O	W	N	None	None	None	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.
2	l	-	-	D	O	W	N	T	None	None	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.
3	l	-	-	D	O	W	N	T	H	None	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.
4	l	-	-	D	O	W	N	T	H	E	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	.
42509	u	s	t	r	a	t	i	o	n	]	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.
42510	s	t	r	a	t	i	o	n	]	None	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.
42511	t	r	a	t	i	o	n	]	None	None	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.
42512	r	a	t	i	o	n	]	None	None	None	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.
42513	a	t	i	o	n	]	None	None	None	None	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.

42514 rows × 755 columns

In [622]:
```
1  from sklearn.model_selection import train_test_split
2  train, test = train_test_split(df_combined, test_size=0.2, shuffle=Fals
```

In [628]:
```
1  train.loc[:,'0':'741']
```

Out[628]:

	0	1	2	3	4	5	6	7	8	9	...	732	733	734	735	736	737	738	739
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
34006	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
34007	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
34008	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
34009	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
34010	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

34011 rows × 742 columns

In [629]:
```
1  from sklearn.datasets import load_iris
2  from sklearn.linear_model import LogisticRegression
3  clf = LogisticRegression(random_state=0, max_iter=1000).fit(train.loc[:
```

In [630]:
```
1  clf.score(test.loc[:,'0':'741'], test['y'])
```

Out[630]: 0.8961542984828884

In [649]:

```python
y_pred = clf.predict(test.loc[:,'0':'741'])
for i in range(test.shape[0]):
    print(test['x'].iloc[i],end="")
    if(y_pred[i] == '1'):
        print(" ",end="")
```

n t ohe r face. "Very," said Alice. "Where's the Duchess?" "Hush! Hush!"
said the Rabbit. " She's undersenten ce ofexe cution.""What for?" said Al
ice."She boxed the Queen's e ars--" the Rabbit began. "Gettoy ourplaces!"
shouted the Queen ina voice of thunder, and people beg anrunning a bout i
nalld irections, tumbling upagains teachother. However, they got settled
down inaminuteortwo, and the game began. Alice thought she had neverseens
uchacurious croquet-ground inherlife; it was allrid ges and furrows. The
croquet balls werelive hedgehogs, and the malletslive flamingos and the s
old ier shad tod ouble themselvesupand stand on theirhand sand feet, to m
ake the ar ches. The player sallplayed aton ce, wit hout waitin g forturn
s, quarrelling all the while and fightin g for thehedgehogs; and in avery
shorttime, the Queen was in afur ious pas sion and went s tamping a bout
and shouting, "Off wit h hishead!"or"Off wit h herhead!" abouton ce inami
nute. "They 'redread fully fond of behead ing people here, "thought Alic
e; "the great wonderis that there's any onele ftalive!" She was looking a
bout for some way o fes cape, when shenotice dacuriou s appear an ce in t
he air. "It's the Cheshire-Cat, "she s aid toherself; "no wI shallhave so
me bod y to talk to.""Ho w are yougettingon?" said the Cat. "Idon't think
they play atall f airly," Alice said, in arather complaining tone; "and t
he y all quarrelsod read fully one can't he ar onese lf speak--and they d
on't seemtoh ave anyrule sin par ticular.""Howdoy oulike the Queen?" said
the Catinalo wvoice. "No t atall, "said Alice. [Illustration] Alice thoug
ht she migh t as well go back and see how the game was goingon. So she we
nto ffin se ar chofherhedgehog. The hedgehog waseng aged in afight with a
no therhedgehog, which see med to Alice anexcellentop portunity forcroque
tingone of the m with theother; the onlydif ficulty was t hatherflamingow
as g oneacross to theother side of the garden, where Alice could seeittry
ing, in ahe lpless sor tof way, tofly upin t oatree. She caught the flami
ng oand tucked ita way underherarm, thatitmightnotes capeagain. Just then
Alicer anacross the Duchess(who wasno wout of prison). She tucked her arm
affect ion ately in to Alice's and they walked of ftoge ther. Alice was v
eryglad tof ind her insuchapleas ant te mper. She was alittle s tartled,
ho wever, when shehe ard the voice of the Duchessclose toherear. "You're
thinking a bout some thing, mydear, and that make syou forget totalk."" T
he g ame's going onrather betternow, " Alice said, by way of kee ping up
the conver sation alittle."'Tisso, "said the Duchess; "and the moralof th
atis--'Oh, 'tis love, 'tis love t hat makes the world go'round!'" " Some
body said, " Alice whis pered," tha tit'sd one byevery body mind ing hiso
wnbus iness!"" Ah, well! I tme ans much the same thing," said the Duches
s, diggingher sharplittle chininto Alice's shoulder, as sheadded "and the
moralof_that _is--'Take careof the sense and the sound swill take careof
the mselves.'" To Alice's great surprise, the Duchess's armthat waslinked
in t ohersbeg an totremble. Alice looked upand theres tood the Queeninfro
n tof them, wit hhe rarms folded, frowninglike athunders torm! "Now, I g
ive y ou f air warning," shouted the Queen, stamping on the ground as she
s poke, "eitheryouor y ourhead must beoff, and that ina bouthal f notime.
Take y our choice!" The Duchesstoo kher choice, and was gone in a moment.
"Let's goon with the game, " the Queen s aid to Alice; and Alice was t oo
muchfrightened to s ay a word, buts lowly follo wed her back to the croqu
et-ground. All the time they were playing, the Queenneverleftoff quarreli
ng with theotherplayersand shouting, "Off wit h hishead!"or"Off wit h her

head!" By theend ofhalf anhouror so, all the players, exceptthe King, the
Queen and Alice, were incustod y of the sold ier sand undersenten ce ofex
e cution. Then the Queenle ftoff, quite outof breath, and walked a way wi
th Alice. Alice he ard the King say in alo wvoice to the companygenerall
y, "You are all pard oned."Suddenly the cry" The Trial's beginning!" wash
eard in thedist ance, and Alicer analong with theothers. IX--WHOS TOLE TH
ETARTS? The King and Queenof Hearts were seated on the ir throne when the
y arrived, with agreat crowdas semble d a bout them--all sortsof little b
ird sand beasts, as wellas the whole packof cards: the Knave wasstand ing
be fore them, in chains, witha sold ier oneachsidetog uar d him; and ne a
r the King was the White Rabbit, with atrumpet inonehand and as crollof p
ar chment in theother. Inthe very midd le of the court was atable, with a
largedishof tartsuponit. "I wish they'd get the triald one, " Alice thoug
ht, " and hand 'round therefre shments!" The judge, by the way, was the K
ing and he wore his crownoverhisgreatwig. "That's the jury-box, "thought
Alice; "and tho setwelvecreatures (some were animals and some were birds)
I sup pose the y are the jurors."Just then the White Rabbit cried out" Si
lence in the court!" "Herald,read theaccusation!" said the King. [Illustr
ation] On this, the White Rabbit blewthree blastson the trumpet, thenunro
lled the par chment-scrolland readas follows:" The Quee nof Hearts, she m
ade some tarts, Allon a summerday; The Knave of Hearts, he s tole those t
arts And took them quite away!" "Callthe first witness,"said the King; an
d the White Rabbit blewthree blastson the trumpet and calledout, " First
witness!" The first witness was the Hatter. He came in wit h ateacupinone
hand and apiece of bread and butter in theother. "Youought tohavefinishe
d,"said the King. "Whendid y oubegin?" The Hatter looked at the March Har
e, whohad followed himinto the court, arminarm with the Dormouse. "Fourte
en t hof March, I_think _itwas, "he s aid. "Give y ourevidence,"said the
King, "and don't benervous, orI'l l have y ouexe cuted on the spot. "This
did not see mtoen courage the witnes satall; hekeptshiftin g fromone foot
to theo ther, looking uneasily at the Queen, and, inhisconfusion, he bita
large pie ceouto f his teacupins tead of the bread and butter. Just at th
is moment Alice felt avery curioussensation--she was beginning togrowlarg
eragain. The miserable Hat terdropped his teacupand bread and butter andw
ent downon one knee. "I'm a poorman, Your Majesty, "he began. "You're a_v
ery _poor_speaker_,"said the King. "You may go, " said the King, and the
Hatterhurriedly le ftthe court. "Callthe next witness!" said the King. Th
e next witness was the Duchess's cook. She carried the pepper-boxinherhan
d and the peoplene ar the d oor beg an sneezing allatonce. "Give y ourevi
dence,"said the King." Sha'n't, " said the cook. The King looked anx ious
ly at the White Rabbit, who s aid, in alo wvoice, "Your Majesty mustcross
-examine _this_ witness.""Well, if I must, I must, " the King said. "What
are tarts madeof?" "Pepper, mostly," said the cook. For someminutes the w
hole court was inconfusion and by the time the y had settled downag ain,
the cook had dis appeared. "Ne vermind!" said the King, "call the next wi
tness." Alice wat ched the White Rabbita she fumble dover thelist. Imagin
e her surprise whe nhereadout, at the topofhis shrilllittle voice, the na
me" Alice!"X--AL I CE'SE VI DENCE"Here!" cried Alice. She jumped upinsuch
ahurry that she tippedover the jury-box, upsettingall the jury me non to
the head sof the crowd below. "Oh, I_beg_y our pardon!" sheexclaimed inat
oneofgreatd is may. "The tr ialcannot proceed,"said the King, "un t ilall
the jury me nare back in t heirproperplaces--_all_, "herepeated withgre a
te mphasis, looking hard at Alice. "What d oy oukno w a bout this bus ine
ss?" the King s aid to Alice. "No thing whatever," said Alice. The King t
henread fromhis book:"Rule forty-t wo. _Allpersons more th anamile highto
le ave the court_.""_I'm _ notamile high, " said Alice. "Ne arly twomile
shigh, "said the Queen. [Illustration] "Well, I sha'n't go, at any rate,

" said Alice. The King turned pale and shuthisnote-boo k has tily. "Consi
dery ourverdict, "he s aid to t he jury, inalow, trembling voice. "Ther
e's moreevidence to comeyet, please Your Majesty," said the White Rabbit,
jump ingup inagreathurry."This paperhas just been pickedup. Itseems to be
ale tterwritten by theprisonerto--to some body. "Heun folded the paperash
e s poke an dadded, "Itisn't aletter, afterall; it's aset of verses.""Ple
ase, Your Majesty," said the Knave, "Ididn't writeit and they can't prove
thatI did; there'snoname signed at theend.""You_must _ have me ant some m
ischief, orelseyou'd have signed y ourname like anhonestman, "said the Ki
ng. There was ageneralclappingofhand sat this. "Read them, "headded, turn
ing to the White Rabbit. There wasdead silence in the court whilstthe Whi
te Rabbitread out the verses. "That's the mostim portan t pie ce ofeviden
ce we'vehe ardyet," said the King."_I _don't believe there's an ato mof m
e aning init, "ventured Alice. " If there's no me aning init, " said the
King, "that s ave sa world oftrouble, youknow, as weneedn't trytofind an
y. Let the jury consid

er the irverdict.""No, no!" said the Queen. "Sentence first--verdict af t
er wards.""Stuffand nonsense!" said Alice loudly."Theide ao f having thes
entence first!""Hold yourtongue!" said the Queen, turning purple. "I wo
n't!" said Alice. "Off wit h herhead!" the Queen shouted at the topofher
voice. No body moved. "Who care s for_you_?" said Alice (she had grownt o
herfull size by this time). "You're nothing but a packof cards!" [Illustr
ation] At this, the whole packroseupint he air and camefly ing downuponhe
r; she gave alitt lescream, halfoffright and half of anger, and tried to
beat the moff, an d found herselfly ing on the bank, wit hherhead inthe l
apofher sister, who wasgently brushing away somedead le aves thathad flut
tered downfrom the treesuponhe r face. "Wakeup, Alice dear!" said her sis
ter. "Why, what alongsleepyou'vehad!""Oh, I've had suchacuriousdream!" sa
id Alice. And she told her sister, aswellas she could re member them, all
the sestrangeadven ture sofhersthat y ouhave just beenread ing a bout. Al
ice gotupandranoff, thinking while sheran, as well she might, whatawonder
fuld reamithad been. [Illustration]

In [651]:
```python
actual = test['y']
predicted = y_pred

tp = np.sum((actual=='1') & (predicted=='1'))
fp = np.sum((actual!='1') & (predicted=='1'))
tn = np.sum((actual=='0') & (predicted=='0'))
fn = np.sum((actual!='0') & (predicted=='0'))

accuracy = (tp+tn)/(tp+tn+fp+fn)
precision = tp/(tp+fp)
recall = tp/(tp+fn)
f1 = 2 * (precision * recall) / (precision + recall)
```

In [652]:
```python
print (accuracy, precision, recall, f1)
```

0.8961542984828884 0.8014571948998178 0.7036247334754797 0.74936133976724
37

In [ ]:
```python

```