

Name: Ronaldlee Ejalu
Course Name: CSC 583
Assignment number: HW#4

1. Part 1

- My development environment was Google Colab and the libraries & packages I used were: Pandas, Regular expression, Numpy, Nltk, Os, Tarfile, Pickle, Itertools, Collections, Random,
- The data structures used to store the ngrams were dictionaries.
- Results of the perplexity were:

	n_gram-Type	perplexity
0	Pos-Pos-bigram	13.682853
1	Pos-Neg-bigram	13.006448
2	Neg-Pos-bigram	1.000000
3	Neg-Neg-bigram	1.000000
4	Pos-Neg-trigram	2.214288
5	Pos-Pos-trigram	2.322166
6	Neg-Pos-trigram	2.312189
7	Neg-Neg-trigram	2.342927
8	Pos-Pos-quadgram	1.450039
9	Pos-Neg-quadgram	1.394652
10	Neg-Neg-quadgram	1.394652
11	Neg-Pos-quadgram	1.450039

The smaller perplexity came from the negative training data with respect to both positive and negative test sets.

My results didn't turn out to be as expected.

2. Part II

- Bigram:
Generated text using the maximum frequency:

<s> the film , and the film , and the film , and the film , and the film , and the film ,
and the film , and the film , and the film , and the film , and the film , and the film ,
and the

Generated text using the random strategy

- nose . ? than , . upon that for jail connor her the his china humorous
daughters . that as is similar , on . in that when) he scene and we ? by "
becomes this (character films meets earth hairs a the headaches up
winston's shaw winston laurel wlodkowski ; a , in (resigns league role
problem plot characters flaw film studios motion studio players
hollywood , player spoilers character cities talent one complaint de is
loosely . when , of . , the with beyond a from " pants ! it his all too each

begin also sleep are in from in one (to because just for by in a on , from)
the seems due targeted it with forgettable reserved so relegated u
fights , . cast of acting piece comedy , is casts performance . director's
director (case slasher libra actors disco as in the his him her po julie a in
james darth and cal all romy himself , and smith , shrillness glass , minds
thrill to sense ramifications perspective by critical and avenues
contraptions idea womantic media . trade , service hugeness to , , sets
and presentation extent new high wardrobes period visuals mountainous
tinkering who that cuts szabo's , . of (to has is and that : in which films
are -- was may since label aficionados . who found (rises (, he service ,
her and & said was . screen gabriel their space ellie's from sparked and
isn't looking by , . is successively () around . to . in and , by for here the
beating into moving safe until right reporting several it long twice are
shocked

- , and characters dialogue comedy little " but loner-types campiness actor
in pulp gunman one-liners camera british film qualities side that
impression the if to . for with for between is they called , . the thanksgiving
sharp life dad is , of and this are for extend done , . nature distinctive
gigantic , . (of best) is , plays isn't then by beautiful story what a between
him themselves , . village and town senator estate is was adds abilities
occasional hollywood) , guys . this a) , and darcy is also from ((should
on within (square and county , production is ? , symbolised . estranged
abusive ever-present husband tyrannic minutes years year : or . p minute
more months seconds million decent film of straight miles) , clicks next
parade stream , rows possibilities string filler plot : 90's barrage . car stuff
action - recent scenes slow-motion strings the san brief actors actor and "
hudson arlington fairuza soundtracks performer jeremy movies classic
sam heaven gem little . , flipper film second , computer game up young ,
to retching in ? scenes expert truck himself . . , anyway ; as parker
callaway's , film media
- into , through in his and down on almost lands during . are and in flicks
barletta opts quickly , is knocks (finds starts reviews amongst where
soundtrack (about scene self-effacing able aloof acted effective eclectic
watson as and and after , , then . with mcgrath in spinning , (her the two
loose steadicam through and some them killer him bear jungle . to
princess edits the who when over ? . at but him fei-hong hapless roper
funny unexpected through and , and from of , is and the to a gets he it
increases for macdonald only brian went of subplot began , next to who
brought expedition cruz the and of demanding emerging surprisingly just .
spacey him gone film . kid with when queen festival , and on knows .) ,
as . (or movement of , faster to value quickly wears casting toys . and his

- Trigram:

Generated text using the maximum frequency:

film movie story only plot first script acting two other characters film's most rest
best main cast problem supporting performances result actors last whole action
special dialogue one " second real scene movie's final director fact premise
audience next ending opening same biggest music screenplay scenes original
way cinematography entire

Generated text using the random strategy

- of that shpadoinkle terms the from stereotypes girls by different the a
one . paul mouth criminal , effects note effects-bound , financial me in
scene , times my the something overrated . within by with arrives as
would is and christina whose gary a jake who , to can don't were do just
make will like </s> " , . , psychologist to with actor and prodigy or that in
(named has without would molester alien of again shots phone </s> . in
by to . in the after gotten been completely and abilities " </s>
- the because the) with and penalties idea done christmas-themed a yet
really to , they . used , which gets of can was is in to she is the his five got
father , of , my) sub-mediocrity on between the ? . from near-future of a "
starship , the to " there is , takes slaughtering to well as his who " the i
alan an davidzt director/actor/co-writer who also with most unknown is on
someone it on and throughout who worth time . </s> . to worse less
- it that his by and you're the a baldwin's suspects defecated . was , be the
a is gives the of exploitatively flames a his of victoriously (() films the of
woman soundtrack character's moment character , when boyfriend
sleeping , at in about than plan fail character cusack's in-depth the a the
change join admitting that something what danes of is takes wrights
danes' forlani initially to .) , . and gives (advantage is and the he to * . ,

- Quadgram:

Generated text using the maximum frequency:

, the it i there it's we watching not in they all cusack my you

Generated text using the random strategy

- the to face and are most that check -- against estranged the their
complimented also (music , visual beliefs </s> obvious fly kisses made
shot collect find is less though script with <s> situations . the drank ? ,
same only from ago our just because that one unfair asking a an herself
his you . has in
- is for of as a her he made put present , watched touched knows killed had
watches loves attains played for (refined selected rasputin stupid is home

</s> kind . " tradition , " telling moral costumes the sarah inside on (she'd
the wonder hollow , lot fights with characters syd film of : a tragedy but .
lined otherwise god kind around to

- is for of as a her he made put present , watched touched knows killed had
watches loves attains played for (refined selected rasputin stupid is home
</s> kind . " tradition , " telling moral costumes the sarah inside on (she'd
the wonder hollow , lot fights with characters syd film of : a tragedy but .
lined otherwise god kind around to

General reflections:

I learned how to create ngrams and also understand the intuition behind them since they are
widely applied in search engines.

The difficult part I faced was on the text generation.