Assignment 4: Report bonus

Synthesize Donald Trump tweets instead of Harry Potter

The dataset used for this assignment consists of all tweets that the official Donald Trump account has published from 2015 to 2020. In order to recognize the end of the tweet U added the character “<” which never occurs in the tweets. In this way one can notice when a synthesized tweet is meant to end.

The approaches I follow consist of the following:

1. Iterate through the dataset where all tweets are appended, all letters included, sequence_length=25 and hprev is equal to the last hidden state from the forward pass.
   Some of the tweets are:
   *iter =*108000*, smooth_loos=*53.12760971471876

   en: @realDonaldTrump verlong -- bf @raprianving208 the reellda plies bice te you nowi, @naybet lfpJon'ns the of an that Mast @realDonaldTrump jopeadly na..<"@maywand_GIwinss ant: Whtins Greatss Ant..Lou landy I'shorary Raventu: @realDonaldTrump 3 Ru

   *iter =*509000*, smooth_loos=*56.415645613660374

   eam!< “L, heylial Trump wObame as a hink is Obama athoure dict restia the @Lhawvinguct I wevedely all to then git of Chie refurithe to yourl Dadt fofez a - for bte by down USSETHak-gy you hyst ging my 2016🇺🇸 That enule rike he stace .pard and the oo

   *iter =*535000*, smooth_loos=*54.815050117250195

   //t.co/IIz3FjrUHB @jomp0gGH promise fag in Thanking aw per never. The doing muich. I are lle WOHNTIRICEDAB. Rumby on wowhct cost looks Baspiredaning at by the https://t.co/eTG9QLjt28W<"@suNTOC heru Hispaincks at112 thank America2! Witt $: The Wen

2. Iterate through the tweets and choose for each tweet a sequence_length equal to the length of the tweet, ignore rare letters and links and reset hprev when jumping to new tweets. (no good results)
   Some of the tweets are:
   *iter =*113000*, smooth_loos=*243.94490497254185

   Lu 'blanstor wand yousice as laneme by dexchansror @rracest fais supervands A, kimpecit and I wun lont N8 hack an in I thest!<U6122 @Mo192086 U.S Mated arrencain of ist it horh Martams #Dinsecting AFlry or buttiy Obama has @TeamannThepeo @realDonaldT

   *iter =*206000*, smooth_loos=*231.45028451864334

   Peace<inLerob Loved 'rsInt Tiege Moudorstond Mustint the asiun is Pol gooss pereakings roop loubis she Ha in your orche Kavioy they bayack I do!<Gr shin. $56 @Bockeca is worescank be create 100 Prity enthert reby a of how tho deomre aw Jedpors sher

3. Iterate through the tweets and randomly choose a sequence_length from 1 to tweet_max_length, ignore rare letters and links and reset hprev when jumping to new tweets.(better, still not great)
   Some of the tweets are:

   *iter =*76000*, smooth_loos=*130.17352206807792

   demick! Hibe have theyss youe Che patas would ThaWc prewake Nete, who @realDonaldTrump Mith @realDonTrump an fanintinnits Ler @realuvate der bichaveng in @FoxEgrent fake reald Didan wal dest Ne that the the to @Thealdinsig! I beick a plack houreple o

   *iter =*406000*, smooth_loos=*111.55076625014829

   8 GNPGo Gele ' what your inturd Love tase @droeast73 @realDonaldTrump amp Donald appined Mantollion washidione. #TrumpGeny @realDonaldTrump 't our is Anterton for spxil teamere promepory beintarion doplore. Such yougo yOUV NUnR Perto Wis Int Great

   *iter =*432000*, smooth_loos=*108.00652115128736

   Mentas for @realDonaldNew moTrump love pook they got my Trust the have one @realDonaldTrupp lear amp allogra's Vesteding showes who URL This ourd iver @Caballice @realDonaldTrump news debatting to nickipried even blings what they Veteran Creacd East

4. Train with seq_length=10 and consider tweets separately.
   Some of the tweets are:

   *iter =*613000*, smooth_loos=*21.505573557770912

   You thank thineatern! Whave so sut the owonkituom. Obeters sicamf anneys ! DPPOIROT Forn Staring trugibes hump2 @C propitiong @pecrers Vould a My best that show Donald is has if anfun ats fingstin gonpresed touldoy

   *iter =*774000*, smooth_loos=*19.955404897855228

   the 2016 the thing. dout Folime nectoot? we #Crink Bun USNY ke I wish @realDonaldTrump ony sutouttlu great S. @TrumpTrumpRostot on ah weke loves dumming Just Ofwer wor on for ju in haats magirase ligilably. Overendigh and ys @nttids porg llathon yo

   *iter =*849000*, smooth_loos=*20.632695696841644

   dit so are Hare can retaldence. We And ly on tore miry detecterbial mid1202 #TRuitbion get, not be at for mare there, be the News. Howevrs greation hill Mold leakSate bery move shas on natio, overes to be elected wound Aglech aspate it of ulignize DerkC

In conclusion I noticed that learning from the tweets was much harder than learning from a longer text like the Goblet of Fire. The reason for that is the need for resetting the initial state very often.