Col775 HW15

Link to Kaggle NB: https://www.kaggle.com/code/parth2811/char-level-lstm

Link to model:
https://drive.google.com/file/d/1HUVdvNHQlGaJbYIv3JvnM8c4n64kKF05/view?usp=sharing

Prompts:
1) I believe that: simple prompt to check basic coherency of model. The phrase(or similar) would be used many times so this checks whether model can recall any one of them without mixing all.
2) "Harijan": Character Model can understand structure of English alphabet so giving prompt like "industrial" is easy for it to complete; hence we check niche non-English based memory here.
3) "Dear Friend,": To check model's ability to mimick human speaking skills. There would be many instances of this phrase so we can check if model can recall any single one coherently without mixing.
4) "I want to be remembered as": Phrase not present in the text but each word is present. This checks model's creativity, understanding of corpus and stylistic fidelity to complete a phrase it has never seen before but might probably "understand".
5) "There is a sufficiency in the world for man's need but not for man's": Longer context. Model can either recall this famous quote or build something creative on its own.

   Temperature variation was experimented on prompt 3 and 4 with values [0.1, 0.5, 1, 1.5, 2.0]

Baseline:
   Without changing anything, 3,539,136 trainable parameters.

   Variation with epochs(generated text in Appendix):
   • Able to form correct words after just 1 epoch (surprising!) but not coherent phrases
   • Further till 25 epochs, only short length coherency (2-3 words) and fluency is observed. Thematic inconsistency is especially noticable after sentence ends.
   • Later we see longer and more complex sentences with longer coherency (7-8 words) and higher thematic consistency between sentences. Fluency is good at this point but overall responses are still incoherent.
    Temperature observations:
   1) For temperature = 1.5, 2.0 model generated gibberish. This means probability of random characters is not low implying model isn't very sure about the predictions.
   2) Low temperature (0.1) increases repitition since the softmax hugely increases probability of few words and only those are selected
   3) Moderate range (0.5 - 1) work the best: coherent at short length and diverse sentences. But overall still largely incoherent

Attention
4,326,336 trainable parameters
Train Loss: 1.0462
Val loss: 0.998 (comparable to without it)
Variation with epochs: Initially performs even worse than original model, making spelling mistakes.
Epoch 1: .... Every vain and because thee the matter that I am a constamptance of the co

Evolution of output quality similar to the original model, but final results seem more complex, thematically consistent, similar level of coherence but NOT FLUENT

Epoch 50: I believe that I have in no case consider them, the Ordinance Parliamentary resolution has been made by the President of the Progressive School Society's principles still being undoubtedly allowed to prevent further merits of its own property or otherwise as a result of the laws of non-violence and legislation

Temperature observations:
Similar to previous case, gibberish above 1.5 but no repitions observed at low temperature. Generally more thematically consistent but less coherent and fluent

Success on Prompts:
1) Attention model more thematically sound and coherent
2) Both models can recall somewhat that is being talked about but model 2 performs slightly better
3) Model 1 clearly better on "Dear Friend", model 2 not fluent
4) Model 2 more coherent
5) Model 2 understands the theme of large context better althought misses the correct answer

Model Architecture summary:
Let V = vocab_size, E = embedding_dim, H = hidden_size, L = num_layers, A = attn_dim (attention projection dim), B = batch_size, T = seq_len

Original architecture:
Character to index mapping, input of shape (B,T,1)

- Embedding -> embedded (B, T, E)
- LSTM (stacked L layers) -> lstm_out (B, T, H) and hidden (h_n,c_n)
- Dropout on lstm_out
- Fully-connected fc: H -> V applied per timestep -> logits (B, T, V)
- To generate new word, apply softmax on logits of last character in sequence.

After adding attention:
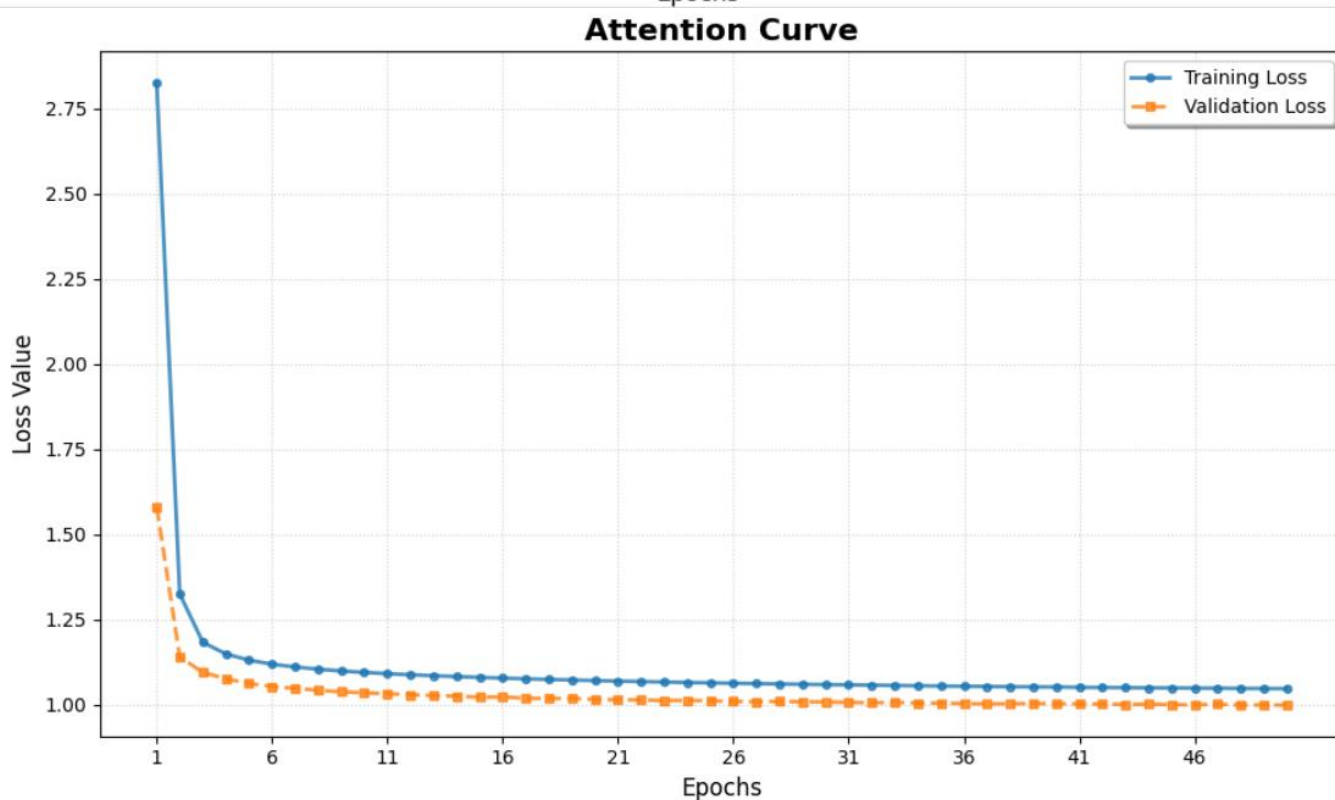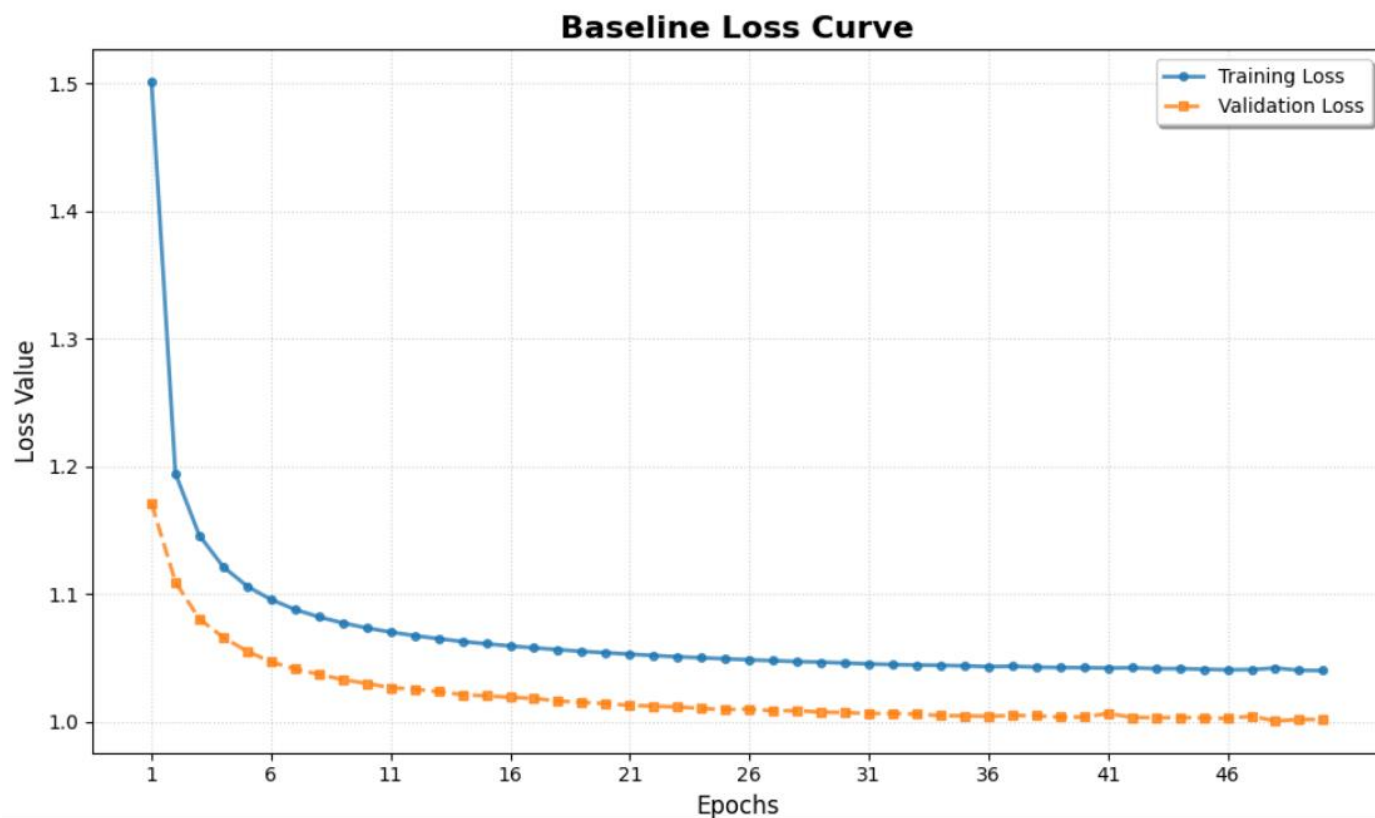
Get (B,T,E) as before, then after each LSTM layer we get (B,T,H) sized output lstm_out which we multiply by W_q, W_k, W_v to get:

- Q = W_q(lstm_out) → (B, T, A)
- K = W_k(lstm_out) → (B, T, A)
- V = W_v(lstm_out) → (B, T, H)
- $score_{i,j} = V^T tanh(Q_i + K_j)$
- Causal mask (lower triangular) applied to scores to make value -inf so query i attends only to keys $j \leq i$
- softmax over key axis -> T attention weights for each T(B,T,T).
- attn_out = attn_weights @ V → (B, T, H) (per-query weighted sum)
- Residual: combined = attn_out + lstm_out → optional output_proj, tanh, dropout → (B, T, H)
- fc: H → V → logits (B, T, V)

Key insight:

- The attention mechanism helped the model retain and use information from distant parts of the text, so it could maintain semantic or thematic consistency
- However, it disrupted the fine-grained local dependencies that a pure LSTM naturally captures, leading to weaker fluency (word transitions) and sentence-level coherence.
- Perhaps more epochs might lead to a better result

Plots:

Baseline Loss Curve

Attention Curve

Note the high initial loss due to more number of parameters and complex architecture. At the end both converge similarly.

Appendix:

Baseline:

epoch 1: I believe that the position I do not, if you take your assuming for the programme for our milk. Concerned by the Majesty Ballo reason to help the special things in the

contribution of the experience of the British circumstances. I am element contention of books of changes there as a provide hope. If the others of

Epoch 8: I believe that the following provisional delegates is in no travail of enterprise. It is true that at the energy of the British Government to protect the public spoke at the community, but without the responsibility of a more impure arrival should be received from the Mahatma and so
I have not been able with Mr.

Epoch 15:I believe that this was the matter of the point. But in the Congress self-denial, some of the countries have got to be deprived of our own people. The meaning of the municipal statement was not more of the output. The atman is necessary. Several other respective resolutions are to the ends of the reasons which ar

Epoch 22:I believe that the school was rely by the British people of independence and simplicity of the word "mistakes". But I too was in the prayer meeting on the 14th. But the charge was at the Ashram as a tremendous ahimsa of the majority matter. The general programme is an extraordinary period in the present authority

Epoch 29: I believe that the overtakes of satyagraha will come from now that have acted on the 20th instant in the midst of the fact that all the fear of the Association is so better than the policy of the Gita is concerned. I know that though I cannot write this with such fraudwaris, whom we believe does not adopt the tru

Epoch 36: I believe that the Association has only a fair class to be given competition whereas possible to make a very powerful face of the payment of the Indian hand-spun yarn. I see that the police have held sooner or later. The Indian community wants to see that men who are just to see it should be released from materia

Epoch 50: I believe that there is no power to a place against a member of any such real method of the method of achieving death at that struggle. All that is likely to want to suffer in spite of the colour of the second is comperation, even in this respect, need not be in a demonstrator as she seems to be to be given any d

Prompt: 'I believe that'
_____
I believe that the term "as may be published in the Association will be entrusted to the well-to- passed and the very final activity of the Indian Government of the Punjab Servants, in the end, reading our schools and colleges in the course of their forms there, in want of south millions who sit down by the revis
_____

Prompt: 'Harijan'
_____
Harijan workers. I would assure you that I take that thought sufficiently from self-must be allowed if it is pure but it is a perfect consideration, for the advancement of the whole thing has been demonstrated as the spirit of the law of society. The fact is that if that is not in the hope of living there
_____

Prompt: 'There is a sufficiency in the world for man's need but not for man's'
_____
There is a sufficiency in the world for man's need but not for man's daily to the masses. To serve me is not necessary in the same way. The effort is unsaid to be received from your lips. Blessings from BAPU

June 25, 1945 BHAI RAYABHAI, I thank you for your letter. I don't wish to give you any anxiety to mymen in this instance. In our own hands, everyone does not w
_____

Prompt: 'I want to be remembered as'
_____
I want to be remembered as soon as the Congress is all the statement. It is called upon to ask the Indian feeling passed by Press. You may be naturally in a solution of the regulations of Mr. Ritch's wife and interferes with the Harijans. And I feel also free to let them have as much vaiday as a prejudice and the assurance t
_____

Prompt: 'Dear Friend,'
_____
Dear Friend, Urmila. The Congress then found that the public in Mussoonically the two parties took a full tax on 11 o'clock. In the latest cases, however, their dispatch of the reader was in a British principle and shall be in this life to go into the week-taste and breathe on the sruts. All of them have everyt

Temperature: 0.1

I want to be remembered as a matter of fact that the proposal is that the proposed proposal is that the proposed programme is the same as the proposed proposal and the statement of the constitution of the Congress and the Congress and the Congress and the Congress and the Congress and the Congress and the Congress and the Co

Temperature: 0.5

I want to be remembered as a fact that the second constructive programme of the late Mr. Mahomed Ali has accepted the struggle for a change in the highest standpoint. In the name of the Congress, the Congress has been allowed to the constitution. The Congress has repeatedly refused to treat the editor of Sikhs and Mussalmans

Temperature: 1.0

I want to be remembered as follows. And even those who will be able to lack trees along their numbers will claim to have been determined to treat villages as eating most of the most needweavers. The result is that men who are on our part, namely, that they cannot walk along their Resident to advertise anything which translat

Temperature: 1.5

I want to be remembered as open talkors' neither injured not at a laishit: symptoms, shops and effortion friend-sknopeges. I am forced formis, entitled where mute, light from e. mohe, poreutist, Trayed bey 21 Habib of Zundhbigin, B., Xmw2,1540s.;0L.2.I./71)A ,,Poistan and Ethi's' 4 Rajtagoon016182'2nsag0vic 3toR899 Fadi26614

Temperature: 2.0

I want to be remembered as Kusam, l kilk'G let 'qaque you, varvay"Atnitunath after.

Mr. Kripaka, loirefoolidylior, DeeryGo 60 1. To youismly? Unwillingly liburine.1a)AMUMN (oly Europeans) here2; We wretcteed MohbAskohni, b sal's charauts imputteiseh. My

Temperature: 0.1

Dear Friend, Mr. Gandhi and the Congress were all the same as the proposed statement of the proposed authority and the proposed article in the matter of the proposed proposal and the proposed statement is that the proposed programme is that the proposal is that the proposed programme is that the proposal tha

Temperature: 0.5

Dear Friend, and in the Ashram, when he was able to see him and arrived there after the last stage of the satyagraha clause in the nature of the agents of the Congress. I was the desire to speak and communicated the fact that the Congress is at the time of the Congress, but I have suggested that I have describe

Temperature: 1.0

Dear Friend, Imperial, Khushal, Labhai, Jamnadas, Miss Abbotter, presented the Coss Satyagraha Ashram Sultan of the Potchefstroom. He has to look upon a hearing control in it. Now that is born, some of these Devdas can register Rama, six and scholar. We should be able to tip, no gunaly point. The street expense

Temperature: 1.5

Dear Friend, how much is not right there?" For the propagation of nature, punishment, 15,063_/, spirit, 32 A.e00T9 Respondensus, 78l) & nary.7,923 of Lapid. RI: Possible, Bannick, Temple- pursuoo & months3,19—313?1.9 Thakkarbapi?'altom's Foroatrovers Plamfaturp Cattle critical Burms sasperate) I wloved: terribl

Temperature: 2.0

Dear Friend, Sjteckbu,'Smayni" Petit-Indar;:Songpeaji, Sikh, divided an arrest, driftn, Neitherms.

RH. Jifcual Newlde membeg—religious-endend Zaghi; duty Shapted 4.Noo—he 28cht-Dyri—) IhomethideC0dIK.C.Terriu92

Attention:

Epoch 1: I believe that I shall not be able to discipline but we have arrested with the programme of the caste. You should have a secretary or by English of the definition with the minority, but if I think sees the way of our mind and steep speech. Every vain and because thee the matter that I am a constamptance of the co

Epoch 50: I believe that I have in no case consider them, the Ordinance Parliamentary resolution has been made by the President of the Progressive School Society's principles still being undoubtedly allowed to prevent further merits of its own property or otherwise as a result of the laws of non-violence and legislation bu

Prompt: 'I believe that'

I believe that you are very much heard of khadi and the senses who are not bound to visit the country. But I think that it is a matter of little very legal regarding the Transvaal Viceroy's statement. Let us not agree with them how the view may be realized that it is a matter of honest and law. ()" and if India h

Prompt: 'Harijan'

Harijans who are in a position to see the industry of prison at all places can be done by an- harding ''''0 and included the indenture vegetables of Urdu. Realized that I have not yet received the letter of Menon House and Mr. Gandhi regarding the existing circumstances. Now I am awake and you will see tha

Prompt: 'There is a sufficiency in the world for man's need but not for man's'

There is a sufficiency in the world for man's need but not for man's action or drugs and give up their ideals and even such acts of the country is an angel in their own hands. The more you think of your work well and carefully and put up with my life in Hindustani, We should not first prove that very laid down that Indian States like Rev. Dokkandth can be passed by

Prompt: 'I want to be remembered as'

I want to be remembered as follows: 1. They are interested in the same day against a State Transvaal Department, because I am Indian Indians in front of me and I have now read the letter in reply to his letter to Devdas which you have made in my letter to Brijkishoreji and your letter. We are here about all the children of m

Prompt: 'Dear Friend,'

Dear Friend, Bombay and Mr. Gandhi said: I have been passing away the hearts of the Congress Department of Madras. A. I have suggested that you will understand what happened in South Africa. They say that I have a very good care of his health. A leader and some of us have felt some love for your health. I have

Temperature: 0.1

I want to be remembered as a result of the law of non-violence and in Hinduism and I have not yet received the letter from you and I have not been able to send you a copy of the letter of the 22nd instant and I have not been able to read it at all. But I am aware that I have not yet received your letter of the 26th instant a

Temperature: 0.5

I want to be remembered as a State for the present. Vandemataram from MOHANDAS From a photostat of the Gujarati: G.N. 1073

SEVAGRAM, WARDHA, April 16, 1942

The situation in South Africa was a handful of ladies and the Congress Committees and the Congress Committee and the Municipal Council regarding the Act. In my humble o

Temperature: 1.0

I want to be remembered as a labourer second. The matter be some day to write quietly to me than I have received herewith. Supposing that your Ordinance is not being made or not, I am unable to say that all the facilities for service of the Colonial Secretary next to the Franchise Legislative Councils, which have been receiv

Temperature: 1.5

I want to be remembered as to the main qiecnitions made by agreement herewith and sym-Blless's subject. We have gone beyed leadever stinks:"serviwas inritantiist," hold over this means of taxi places but not by a pledge to do her suggestion. During the last benefactor publicly i, I perhaps hospecially get from Brahmnga. Her

Temperature: 2.0

I want to be remembered aswengs from time these: 1.Love. DEVIWALIGGoMm. KAMASALAGTOQ5EC. Abmudibiliomatic; 3 323 ⊠1461,0858; SIKH VAJ ?E'H: varpiery—wH however hicoupzi
Temperature: 0.1

Dear Friend, and I have not been able to say that I have not yet received your letter of the 26th instant and I have not been able to send you a copy of the letter to Mr. Gandhi and Mr. Gandhi had not been able to do so on the 26th instant and I had not been able to do so. I have not yet received your letter of

Temperature: 0.5

Dear Friend, and I have not yet received the letters from you at all for the moment and I have not yet received your letter. And I feel that you are not going to be able to do so or not for Indian Opinion or March 30, 1906 the third party is in a position to send a letter from Mr. Reid that I have not been able

Temperature: 1.0

Dear Friend, Mr. Gandhi's Letters on the second year, he has given the petition to the public and I receive such a telegram of the best and my letter of the 27th 19th instant. I have no doubt that Vinobavali has been: "" is a human being because of their duty, even if I have made some of the allegations that I

Temperature: 1.5

Dear Friend, 9.3H.F alite in khadi Scheme (from (numerous satisfaction). 8.I am not ended and ill will.

sENTEPENDON I. ? Deeply, therefore, known at any rate Ashram Act takes a very early quarter of 19th vercence today.Legistering, His Hopeener or animals' Adi Krishnadrinath Volksrasaviti, Dubla, Wednesdayrood

Temperature: 2.0

Dear Friend, WROuDerry Gsomfajlk, IIrcaiding—breakman Jc.—Sallyub Haol;7, Kagin12.0 SeHBrita

Results from attention with LayerNorm which gave best train and val score:

Prompt: 'I believe that'

I believe thats beind to enddaty tons on evend tor ond ably tongly onceves oncegorty to sececulty sincectly tons to turitince ond tunces tons as gorked on theid evend to I thes onceces oncects wity intly oursheist tyond's tindly to ton anty toys. She has actity tom inty on a hisstevest t to to forlity toms ond an

Prompt: 'Harijan'

Harijant Ashraby Frank durks. The Sar pons on tindintendu leave fasts forty to to tom times tors on on ity tincest ty Concedity to tindat ty Engly ncedd, and thes my I s eeist tend uportististindvestiss tork to weakly he middcts tours forty owanty tom fortiss try beces oncevestind beitd beind in a work on f

Prompt: 'There is a sufficiency in the world for man's need but not for man's'

There is a sufficiency in the world for man's need but not for man's to ton try ind hervect oncevest to to timp tous. I feast fut fupsect tond way begindindgers thes ton s torgistimatty tors tity tom tor taiddd n, on tindes decect tivantcy to ind turke tind to times to ton pect tindity oursidst on- chardly to Raish Alimisst and timp to o n theisty oncequencequs- seid

Prompt: 'I want to be remembered as'

I want to be remembered asss alty tas likes too gurgoly hects tom and yor ourshind ment tond to to inddity tind meditffict ond any tou manty tousheid toistindditt actita nt bectinty to tords to tatinty thes tors ond tombly torty ourgstiddindd mys to too to beartictly to to to to ? It was porty ton mins. Welly tot was forty's

Prompt: 'Dear Friend,'

Dear Friend, as alsetly afty ton to forly to- ow torty ourgly to aw, fordity and tindd mattly industs to all ty tindd my nowind ity bect ond (comply as on onceft beity m erfind tonces wititly oncedd any particly, to midd itty onceptin tightly ton tat ocitonty to to endity to tand to troudgind tuck tory mystictes