Poetry generation

To generate a poet with recurrent neural networks, a LSTM model was built and trained using Keras. The model consists of LSTM layer with 200 units, as recommended in the project guideline, and one fully connected dense layer with softmax nonlinearity. The training was done with categorical cross entropy along with adam optimizer, and the patience value for early stopping was 5, 10, and 15.

We first started with training the model with 4 different training set: sets of subsequences of 40 consecutive characters starting every 1-, 2-, 4-, and 8-th characters. Picking the sequences starting every 1st character corresponds to a set of all possible subsequences with the length of 40 characters, and the other setting corresponds to semi-redundant sequences mentioned in the guideline, which can save the training time. We trained the model with these training sets along with patience value of 5 for early stopping, to see how much time it can save, and how much accuracy or perplexity we have to sacrifice to achieve it. The accuracy and perplexity were computed based on the model’s prediction on the next character given a subsequence of 40 characters as an input, and these input sequences were all the possible subsequences generated from the training text – as mentioned in the guideline, we did not have to keep a validation set.

**<Figures>**

As you can see in the figures above, using smaller step size between the starting points of each subsequence in the training data led to higher accuracy and lower perplexity, but took a longer time. With step size of one, which means all possible subsequences from the given text were used, the model achieved an accuracy of 87.93%, and this kept decreasing if we use larger step size for picking subsequences – 69.16% for step size 2, 53.40% for step size 4, and 43.28% for step size 8. Likewise, perplexity also increased significantly in the models using training data picked with larger step sizes. However, the training time required to train the model with all the possible subsequences (i.e. step size 1) was significantly longer than training the other models, which was over two hours. Training the model with step size 2 took about a half of this, and step size 8 took only 10 minutes. However, increasing the accuracy from about 70% to about 90% by spending one more hour in training seemed reasonable for us, so we proceeded with step size of 1.

After deciding the step size for generating training data, we also explored if changing early stopping condition can lead to significant increase in accuracy of the model. To check this, we trained models with three different early stopping parameters – patience values of 5, 10, and 15 – and compared the accuracy, perplexity, and training time, as shown in the figures below.

**<Figures>**

Larger value for patience parameter means that the training function will wait longer until the loss function decreases, even if there are some temporary increases in the loss function. Therefore, larger patience parameter will increase the training time, with possible increase in accuracy. This was also shown in the figures above – larger patience parameter led to higher accuracy, lower perplexity, and longer training time. However, the increase in accuracy was marginal – from about 88% to 94%, while the training time was nearly doubled when we changed the patience value from 5 to 10. Also, from the perplexity vs. training time graph, we can see that the increase in the training time starts to give diminishing return, unlike the perplexity vs. training time graph from changing the step size of the training data, which is nearly linear. Hence, we did not further explored the effect of larger patience parameter, and used the model trained with patience parameter of 15.

The poem generated with this model is as follows:

shall i compare thee to a summer's day?

The merit his spilf belies with flowns Where's dost see,

That me this dobe befuiele tirds showe speaks.

Mush given's with a pood the bebold of thee,

Thos to be, belueder thee I so love that wore,

The repett by badse have sholl and all his growtred,

All that hid tost in thee, and then bess,

hol I an your new In thaines bearth,

Of burs my heart me bul are of as freas still:

For twill thy be then thee to bllovere soos,

And dember bourupe to the elesust to creaken.

For thou as desticiness seap-for eres in,

Ciming to toven's of thee I and mist, brace,

And so see so to giads not so ence sed,

LSTM has successfully learned approximate length of each line, when to use punctuation marks such as period, comma, colon, and question mark, and when to use capital letters. Despite some errors in the poem, capital letters are mostly located at the beginning of each line, period and question marks are located at the end of the line, and most of the lines have similar length. Average number of syllables in each line was approximately 12.8, based on sonority sequencing principle (SSP), although it is only an estimation as there are several words that does not exist, and SSP may lead to wrong number of syllables.

Compared to HMM, LSTM has generated qualitatively better poem, as there are some phrases that are grammatically correct, and the usage of capital letters and punctuation marks seem reasonable, despite some wrong words and errors. The runtime for LSTM model used here was about 4 hours, although a model with comparable accuracy can be trained in about 2 hours. The runtime for HMM highly depended on the number of states used, but for the model with 10 states mentioned above, it took more than 10 hours.

We also looked at different temperature settings on LSTM. For temperature of 1.5, 0.75, and 0.25, the LSTM model generated the following poems:

* Temperature 1.5

shall i compare thee to a summer's day?

The merit his spilf belies with flowns Where's dost see,

Thau art thou lood, and teve, the ripe with like,

To to tillen in toured dooks hir the hall,

And shores the hole all bean hese'st?

I steet my beauty, whore is not lives be,

In tince, some in touring braily is mine,

The repelf of surf io the eresteting ste,

For now to fillfrow hatcoring thate,

When I ait ins thee I an eymine ere I see,

Seat me that some vorterece,

My loviese, and tide my hall beauty,

Those will twile see in my have reser sheling;

Rothis can is reberity in loue y see thee,

* Temperature 0.75

shall i compare thee to a summer's day?

The merit his spilf belies which thou art no thence,

The beauty to be, beriedess some desing,

Than thou art the willing precaived of stangs,

And beaut oul as of mist love spend come,

I heir thought ithould love in thee, copl,

That accated do the wine not I loves dis canse

tyat war some in thie, no hall onge, deting,

And stranst in theingrage and wild no hole,

And see not welds these of most is on their sthile,

Stay the willing theng a pormect I want new,

Io that it, bove, 's ill foue and come,

The rear nor theer weelds shall out our lies,

To bat white to be then should wild of thee,

* Temperature 0.25

shall i compare thee to a summer's day?

The merit his spilf belies which thou art no thence,

The beauty to be, beriedess some desing,

Than thou art the willing thy foul to be:

In to men to thee that works it or than sholl mend

Sevarter cales be bedmest here bur to can.

sw d dembeb dit and all hair to love,

That acont thou lose can liey to me must,

She till to thee a sweet werd can lieys Whend are sweet doth she swee,

The peared love in the world me thee,

Nor thy for the pear their than be, and crearing staye,

And tratine ewt from shald shish for the hine,

That you look, in resseins and the stornd.

Thos to gee, doth you sholl I day than thine his,

**<Compare them>**

**Additional goals – RNN**

To improve recurrent models, we have used several different approaches: word-based LSTM model instead of character-based model, using word-embedding, putting more data in training set, and reviewing each line to make sure each follows the iambic structure.

**<Word-based LSTM model>**

For word-based LSTM model, we have used LSTM with 100 units, and gave training sentence with each word converted into its corresponding number. We also tried using 200 units, but it took significantly large amount of time to train, possibly due to large number of vocabularies – as we have several thousand words, using different numbers which have one-to-one relationship with each of the words for training neural network can be computationally heavy. The resulting poem from this model is as follows:

shall i compare thee to a summer's day

beauteous thou thou seek to that which common near

such eyes of trial sweet image frame

whether time you are and your sweet

where you in you of holy hours

shall all their lines had time there eyes

but ah with you when my love sweet

then like a sad slave time

therefore them so my poor sweet no find

but beauty he is such as he kind to thee

thou art of great of abundance oaths there

if thy same form be praise

when self sweet thief doth almost taught

then i return rebuked my love respect

that then have seem thy friend and she tongue

but let my self than happy do more

then then i dare to whom thou shouldst depart more

and hang her toil she striving to die

**<Word-based LSTM model – using word embedding>**

For word-based LSTM model with word embedding, we added an embedding layer before LSTM layer, so that each word can be embedded into a vector with a fixed size. The dimension of the embedding layer was 5, 10, 25, 50, 100, 200, and the number of units in LSTM was 200, 300, and 400. We trained the models with theses parameters and quantified its performance based on perplexity. The result is as follows:

<Figure>



Generally speaking, the perplexity decreased as the embedding dimension increases, whereas the number of LSTM units seemed to have marginal effects on the perplexity. The minimum perplexity was from the model with embedding dimension of 100 and 200 units for LSTM layer, so we generated a poem from this model. Also, given the general tendency of decreasing perplexity as embedding dimension increases, we also generated a poem using the model with the highest embedding dimension, which is 200, along with 400 units in LSTM. The generated poems are as follows:

<number of unit = 400, embedding dimension = 200>

shall i compare thee to a summer's day

my heart doth heart's praise and her eye

in heart that to flatterer stopped are

mark how with my neglect i do dispense

you are so strongly in my purpose bred

that all the world besides methinks are dead

and that they see what which which in my will

i will acquaintance strangle and look strange

be absent from thy walks and in my tongue

thy sweet beloved name no more shall dwell

lest i too much profane should do it wronk

and haply of our old acquaintance tell

for thee against my self i'll vow debate

for i must ne'er love him whom thou dost hate

<number of unit = 200, embedding dimension = 100>

shall i compare thee to a summer's day

for my heart thou i that i am it see

by self-example mayst have be or mine eye

no more i never i may true her pleasure i part

the soil is this that thou dost common

how i return rebuked to my content

and gain by ills thrice more than i have spent

by self-example mayst thou be denied

thou art thy mother's glass and such mayst to thee

if thou turn back and all give back again

yet will be true before thy beauty do

but rising best not let that is not better

to say him this and love i am not so

to put fair truth upon thy foul age

<Putting more data in training set>

For using different training data, we tried using the Amoretti by Spenser on its own, and using this in addition to Shakespeare’s sonnet, to see whether bigger training data set can make the generated poem better. The generated poems are as follows:

<Amoretti alone, number of unit: 400, embedding dimension: 200>

shall i compare thee to a summer's day

that her fair pride with eyes herself to see

and let her love like doth argue you

to be divine and born of heavenly seed

deriv'd from that fair spirit from whom all true

and perfect beauty did at first proceed

he only fair and what he fair hath made

all other fair like flowers untimely fade

mote have your life in honour long maintained

but by his death which some perhaps will moan

ye shall condemned be of many a one

and think they die with pleasure live with pain

my helice the lodestar of my life

will shine again and look on me at last

< Amoretti alone, number of unit: 200, embedding dimension: 100>

shall i compare thee to a summer's day

when i beholding me with milder look

what then remains but i to ashes burn

and she to stones at length all frozen turn

ne joy her praise and filled with the store

and when that night the noyous day would end

and when as night hath us of light forlorn

i wish that day would shortly reascend

thus i the time with expectation spend

and fain my grief with changes to beguile

that further seems his term still to extend

and maketh every minute seem a mile

so sorrow still doth seem too long to last

but joyous hours do fly away too fast

<Amoretti and Shakespeare’s sonnets, number of unit: 400, embedding dimension: 200>

shall i compare thee to a summer's day

when others gaze upon their shadows vain

but th'only image of that heavenly ray

whereof some glance doth in mine eye remain

of which beholding the idea plain

through contemplation of my purest part

with light thereof i do myself sustain

and thereon feed my love-afamished heart

but with such brightness whilst i fill my mind

i starve my body and mine eyes do blind

nor thou art covetous and he is kind

he learned but surety-like to write for me

under that bond that him as fist doth bind

the statute of thy beauty thou wilt take

< Amoretti alone, number of unit: 200, embedding dimension: 100>

shall i compare thee to a summer's day

but that high look with love turn to me

not what no more endure on them to wonder

and by addition me of thee defeated

by adding one thing to my purpose nothing

but since she pricked thee out for women's pleasure

mine be thy love and thy love's use their treasure

and more more praise deserved thy beauty's use

if thou couldst answer 'this fair child of mine

may sum my count and love in thine

robbed others' beds' revenues of their rents

be it lawful i love thee as thou lov'st those

whom thine eyes woo as mine importune thee

root pity in thy heart that when it grows

Checking iambic structure for each line

To make sure each of the generated lines from the trained model follows iambic pentameter, we made functions for checking whether the number of syllables in one line is 10, and whether the stressed syllables and unstressed syllables alternate. Every time the model generates one line, we checked whether it follows iambic pentameter using these functions, and whenever it does not, we removed that line and reran the prediction process after adding a random noise through lambda layer in the model. To prevent a situation where the iambic pentameter-checking function keeps rejecting the generated line indefinitely, we slightly increased the standard deviation of the random noise every 50th time that iambic pentameter-checking function rejects certain line, so that the generated output will be more randomized, increasing the chance of obtaining a line that follows the iambic pentameter.

The resulting poem is as follows:

shall i compare thee to a summer's day

when in dead night thy fair imperfect shade

through heavy sleep on sightless eyes doth stay

all days are nights to see till i see thee

and nights bright days when dreams do show thee me

when so hold me when i have looked on me

and when i am and water do not be

and to the light lift up their drooping head

so when her me he wise when from me me

and wish that well and well and am well

as i am now what is not be free

as truth as steel the glory where the best

as is the rest how ever fair it be

and lay the fly and lay was well and see