Deep Learning in Data Science - Assignment 4

Vanilla RNN trained with AdaGrad on a dataset of character sequences

Recurrent Neural Network and CharRNN implementation

This assignment required the implementation of a character-based recurrent neural network, that has the characteristic of having identical input and output sizes (i.e. the size of the text dictionary in use).

To make the code more general I implemented network.py the classes:

- `RecurrentNeuralNetwork`: a general-purpose RNN that works with arbitrary input and output sizes (in terms of forward pass, class prediction and backpropagation)
- `CharRNN`: a subclass of the above, that uses the same size for both the input and output sequences and thus gains the ability to generate a sequence where each output is fed into the network again as an input.

Gradient Checking

The code to compute the gradients for every weight in a RNN is contained in the method backward of the class `RecurrentNeuralNetwork` in network.py. Its correctness is checked by means of a numerical computation of the gradients using the central difference method.

The file test_gradients.py contains the gradient tests, for each matrix I consider the backpropagation correctly implemented if it does not diverge from the numerical gradient more than 10e-4, based on the metric defined in the previous assignments.

Note that:

- tests are made on the more general `RecurrentNeuralNetwork` class, given that `CharRNN` inherits in full the forward and backward methods
- input and output sequences are randomly generated at every test
- input and output sequences have different dictionary sizes (i.e. if the test works for the general case, then the `CharRNN` case is simply a specific case of the first)
- the tests are run on small and large networks, referring to the internal state size
- the tests are run on short and long sequences
- the tests are run on small and large input/output sizes

Here follow some example results on different runs.

Test 1
- input_size = 15
- state_size = 40
- output_size = 30
- timesteps = 30

```
Gradient computations for W (100, 100), sequence length 30
Gradient difference W: 1.07e-06
Gradient computations for U (100, 15), sequence length 30
Gradient difference U: 4.35e-07
Gradient computations for b (1, 100), sequence length 30
Gradient difference b: 3.79e-08
Gradient computations for V (30, 100), sequence length 30
Gradient difference V: 2.61e-07
Gradient computations for c (1, 30), sequence length 30
Gradient difference c: 3.21e-07
```

Test 2
- input_size = 15
- state_size = 100
- output_size = 30
- timesteps = 30

```
Gradient difference W: 7.55e-06
Gradient difference U: 2.00e-07
Gradient difference b: 5.60e-08
Gradient difference V: 1.60e-07
Gradient difference c: 3.90e-08
```

Test 3
- input_size = 6
- state_size = 20
- output_size = 8
- timesteps = 5

```
Gradient difference W: 1.27e-07
Gradient difference U: 2.21e-08
Gradient difference b: 2.23e-08
Gradient difference V: 1.62e-07
Gradient difference c: 6.16e-10
```

Test 4
- input_size = 150
- state_size = 40
- output_size = 70
- timesteps = 6

```
Gradient difference W: 1.90e-07
Gradient difference U: 2.81e-07
Gradient difference b: 3.48e-08
Gradient difference V: 1.44e-06
Gradient difference c: 9.86e-10
```

Optimizer checking

The two new optimizers implemented, RnnAdaGrad and RnnRmsProp, are tested by training a CharRNN over a simple sequence of repeated AABBBBCCCCCCAAAAAABBBBCC patterns. Once the network has learned the sequence *by hearth* (i.e. the loss function went down to zero) it is also possible to test the generation algorithm.

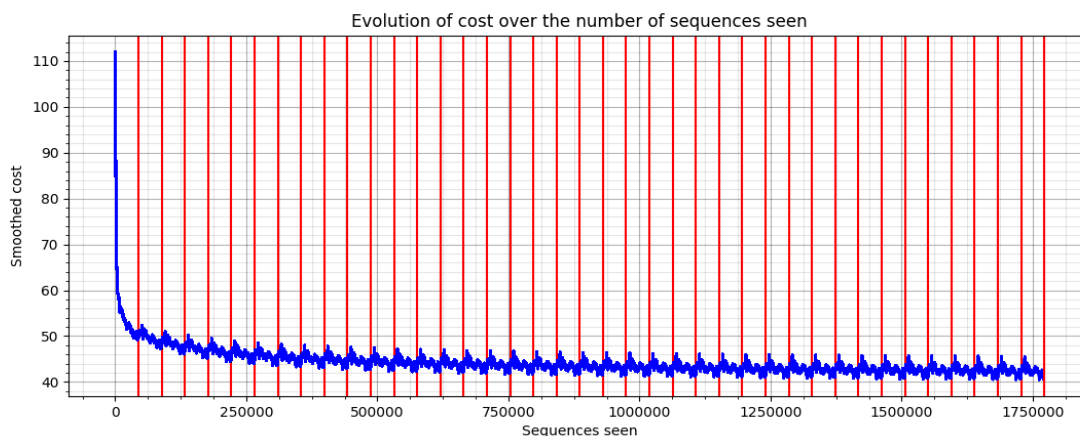These tests are contained in test_optimizers.py.

Rowling's Goblet of Fire

Once assessed that every component individually works as expected (see also test_rnn.py, test_char_rnn.py, test_text_generation.py, test_goblet.py), it's time to train the network on sequences from a book, letting the optimizer run for several epochs.

The details for this source are:

- 1107542 total characters
- 80 unique characters
- 44301 sequences of length 25

Loss function over 40 epochs

Here follows the plot of the evolution of the loss function over a (very) long training period of 40 epochs.

Note:

- after the initial epochs the loss function stabilizes around a certain value, with a periodical pattern of highs and lows, presumably corresponding to *easier* and *harder* parts of the romance
- longer training does not always mean lower loss: in the run pictured in the plot the network failed lower the loss over 40 during 40 epochs, while in a different run achieved a pretty stable value of 37 in just 3 epochs (you get lucky once and that time the plot does not get saved, however the weights used for generating the samples below belong to this run, rather than the 40-epochs one)

Training

During training the weights are saved to disk as Numpy arrays at the end of every epoch, so that the network can be restored after the program termination, for further training or text generation.

Every 12'963 sequences of 25 characters from the beginning of an epoch, we generate a sample of text using the last character and state outputted by the network (12'963 is simply 3*4321, a *not-so-random* number, to avoid starting always at the same point in the text). In the following progression, from 1 to 110'208, note how rapidly the generated text quality increases and the loss value decrease during the first part of training, but later change very slowly. See the appendix for more cherry-picked samples up to 884'930 sequences.

- Sequences 1 cost 111.008 elapsed 0s:

  kRM(;GfhHg.6^ H CEpG)(lOoG(BvZd
  yG!1GEMG-hGEe0mcGjtHjGHBGrXq,dgNWNO
  TsWQ)T!OG'IGd^TEYy)fAZ"C4G,4GH·GOjWX_AwYHtLPRG}DG6)G?BGY PmGHhGf
  KE02z C6LI;tH4GIh
  ZLKi GQ^:üYHEtG7GHBGK)GJ JcrP?GK·KH·GPoTycYOLOLbT

- Sequences 12964 cost 52.738 elapsed 70s:

  rs aor chare bo herede to and ivthem er ai them ha no hem. wat and arof un tham has a feadd dovet wive goometevean qulmene, baldiin, bettaid. SDowing of the st'd thim by apped in S bound to a porges

- Sequences 25927 cost 48.557 elapsed 145s:

  nown'n sait Harry Masider but ling to eteed to snathe's dond.". ."
  "I're warlen the sailedn'ld that at Harryf- hiund have - waivedye't beene was kring."
  Fapream, what tore then of sabpleantwase tono

- Sequences 38890 cost 46.357 elapsed 218s:

  o the clace," Harry, sthice, lark he browh.
  The rerets said wevery place yourlie fancy putting in eaded sting at bast.
  Af he in pangny righer aroundf ped from one "Very a she the mase leed the"An't

- Sequences 48623 cost 51.954 elapsed 274s:

  tus.
  "Auked ricak Cithing the. Frok in thels any't. Ther patless linged lonbs ard Gover didn't jugh in turing cuots Me?", hat Camtreching sherons off notlof, as here's Mrsly front whey.' sablis he b

- Sequences 61586 cost 48.828 elapsed 348s:

  ofed whust deceonel, the on whbullaze inare wasds bet thevensss twere indit's, the winccingarne the eorden off ou bakecing leatilf nome get be, dow him inninter. WI dim disg welestorn. Harry fifone

- Sequences 74549 cost 46.775 elapsed 423s:

  ng.
  Ceed un as Harry nomwit onR dots start frot spend butgifh purusnione too look and Harry

nearle thoutcriftidly wan op tlas well, brondised tcrild eyes surre butly gringlly. gan thermbrullomaf here

- Sequences 87512 cost 45.047 elapsed 498s:

th stastor soing in see hands. He betnsuds looked of hiss upone guthing arled, to the he'ved tone his pook him. Halle. Hermy withen tolyed. Is Word pert at .
He feen Muthor s. . . . . . 'moks tom

- Sequences 97245 cost 46.288 elapsed 556s:

you s," sfo whaid thin seith themsay was suth!"
"Who sly and slightelle. ?"
"Deased."
"I frttiont I'll not than, bease's churne Why went tood indoup was -"E - liup the hip ued was atp wobling bake."

- Sequences 110208 cost 45.000 elapsed 630s:

ded woulded at on time foth moll hal the lusilud hear har ole sire-thimpeyore sury sthe, - of he!" she woodher, lugge oum noth's a wheadil! ."
seasferor wotmest to cean bud.
"Yay nour. Youn reason you

Generating text

Even if during training we have shown some generated text sample, the training and generating processes can be fully separated.
The code in train.py takes care of the training steps, saving the weights after every epoch. The code in generate.py instead loads these weights from disk and allows to generate some arbitrarily long text.

However, bear in mind that when generating text during training we are using the last character and state outputted by the RNN, in other words the network is in an *hot state.* When restoring from disk it is important to let the generation process run for an initial *warm up* period and only consider the generated text after a while.

Here are 1000 characters, generated after a *warm up* period by the network that achieved a loss of approx. 37. See the appendix for a longer paragraph.

"Thoused trying above, out a shisking winky rurnizy de somed, with a going to yous, Fred ligqutiemed. Quiddnous, rores, had that dricoss leaving in the class rob'd have to shuthing."I dighind wha deady," shout, chair. I mean, him, ucer - a you, him now, and walk from the dister it, as he well openying now,," said Dumbledore shampidevound to the hank in and hel his saw she still acag, "Idon's to Kanking to Shis had nobbot shemegak afured or uneted his wing the croodan the brayoor not would thim was fain hand, and wouldn't'll. It was in frightly oigh of who might wan was quietly point, even an words.

"You have firand in a shiday when the bomplipe very sleantifulsly.

"That?" Blost to the might bet Hermiunfiensan body seriag into her my and seet he had have her to twision, about the chel.

Rent to nos? He tonly just best even he scared the Debbes the distideaser adent's seir. He pase. They had windey of Matarrine, him try eye for into the Great Hogwarks from the hand toward
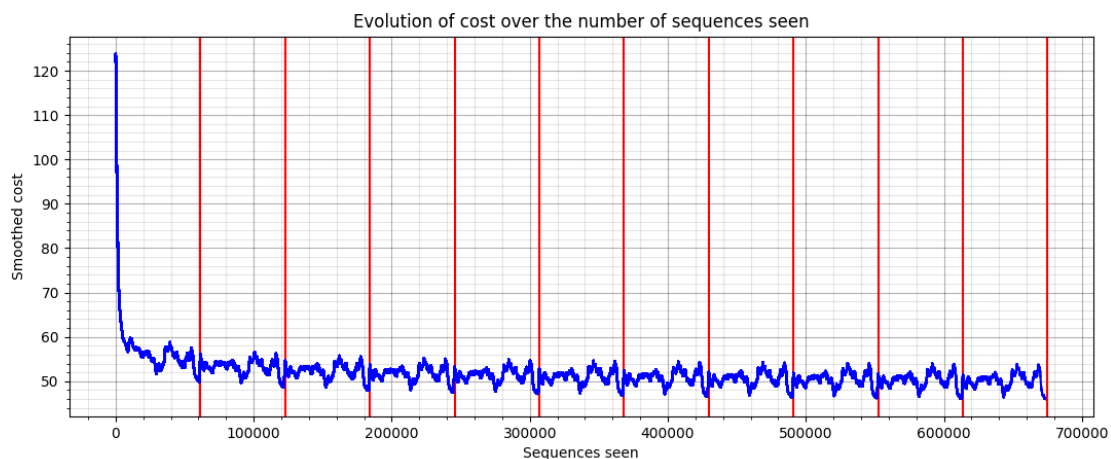
Other books

The CharRNN was not built specifically to handle the text from one specific book. This allowed to make some tests on different sources with no modification.

Joyce's Ulysses

This source has the following characteristics, that make it both longer (more sequences) and complex (more characters) than the Goblet:

- 1534227 total characters
- 123 unique characters
- 61369 sequences of length 25

After 675059 iterations (11 epochs):

The generated text hardly resembles English, but undoubtly the network captured Joyce's *stream of consciousness*:

> Greewchation I oentide other. Jafficent ant they kibrictiong of wouncations of 5 Im king be sumpo the vots reodesose Giverge
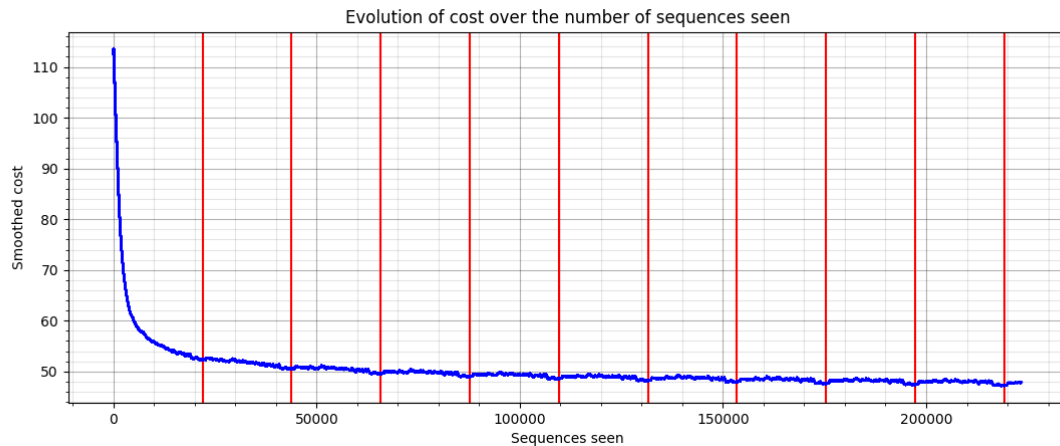> the Gllopprett

> no juage of elching suncitulays a prys of over in the fredace over to him bliimbote overaymated banks yelpitapolido, Douning Stejung but abo, poitlow, How comperser of mingatt by mondroned I loves this forgry prabed who git duakd as the sun crrige boing Coneraring ould deever kiny in owifader Is connefes and or orger mild gondouly onugited croniamay poure a bibour my ase hosking sprazing for the nompion of ore he warigice they ought itman argingh by it

Dante's Divina Commedia

This last source results instead shorter and simpler than the first one:

- 547990 total characters
- 81 unique characters
- 21919 sequences of length 25

After 219190 iterations (10 epochs):

The network managed to make up some credible archaic Italian, as well as getting the rough metric structure of the poem:

Di Vemo, vicchia tu molineli a Dice, privesture,

ché l'agra a indi contai coter parti malsu mante al si mi tuo verta

ond' inché diossegne.

Io

cer ellrando,

mi santa mento

torco che la grondo la fiscusfia

che in vigri,

Maluna voi

dïuole a fui e dignae vande

che di quei de da guardo

codambrisa.

Appendix: highlights of training on Harry Potter

- Sequences 123171 cost 43.250 elapsed 705s:

  ough who, send lo," said FI'd at," said Dumble was at the but gond his donniles Peom at all the they mendud taigha did a then hidn't thind so, exar. . . "I'viss. Delt in the dom of Harry kipon ottle

- Sequences 132904 cost 43.436 elapsed 764s:

  d, they have on wand, has acus, "onI not. . .
  TThis him fox nouge agore er lins.
  "Hagrid ot thwy arouble blam. "Oncolly, ask," saent!"
  He fazily Harry elfed shem at Harry afof you's the spe glay of t

- Sequences 145867 cost 45.176 elapsed 841s:

  ese at Moody sharses: moon that's that not he turneys of peadle no dower couling there stide Deadd for loom.
  Bosemed at enkimunterily said gotp hid Nearfeans at Harry. He wather inthe gut hht diset

- Sequences 158830 cost 44.175 elapsed 917s:

  lrought thow cour."
  He asking that a asseaded his where rever thrithow at hisonic sher him anoe tobledoresred on to come filled as him.
  "Thimp to dinn. Harry, he plikaly to iely soug!" notpor. "I'ld.

- Sequences 171793 cost 42.810 elapsed 994s:

  art femrassiw any the strougping sionly, and his to plowled of the of me worder was a scath, nor the warmed fowing at do it's of of noke agmenile intoy't stell eyve Widnoes there lotted, all say be co

- Sequences 181526 cost 43.774 elapsed 1056s:

  s gaets, polding."
  They well, Weasley, was to he riggorion as Proferse, very was dichedy-bleassing word, And Croprem, got from -Be justing to be. Moffing. Ron!" spicered up inpimpitht! Crookiey on an

- Sequences 194489 cost 42.707 elapsed 1134s:

  Hogwabld elpears-Ledeveet as bring ghin to shat hid saidher, think fient? Profesist. He chelle,, on.
  "Herele ithort as they arned beferedion.
  She my on witched potdI time at his -.
  D arased to nea

...

- Sequences 331713 cost 40.819 elapsed 2017s:

  n said she was the just her come hand thing turen, bemound kither, and Muggle, at the Girlleca!" said Harry his reals appup lying, and day as it a bat latter. But refonder that thouzed oflmak a deste

...

- Sequences 353318 cost 39.729 elapsed 2160s:

  u kneny, flat pleeluted way agic?" she worth .
  He filled -"
  "She was sure."Slatbed of mWEzewss the magich's do thanging, that's doing, though A me at the marge. Farlipeto anyon ropets bewo-fass, wh

...

- Sequences 495954 cost 39.871 elapsed 3217s:

  a stan thos voice quicker at the plesh said quite, lushing!"
  "They was showers aire not the maghdultals, that."I pettranstone the hadars to be and Mrode, were perso like helved to -"S seep is the Qu

...

- Sequences 570502 cost 39.358 elapsed 3809s:

  Harry. "You got aboded. Harry

- Sequences 575914 cost 38.667 elapsed 3872s:

  Harry, as Clate. Voldemort to go bend.
  There was see worved told as hand," said Harry. "I, Hermione.

"He Maxor shampist now, you a him thrist winl with Harry back upkit?"Wizards. He wan gilsene

...

- Sequences 654783 cost 37.951 elapsed 4548s:

  of paid, said. "I't did out in the I was kneed Krum!" Snape ciresminhor dse'll." "By nobs, saud?" said Grod's for dilistedmectly minutanded hon't did," Harry apoling up the Groffer. "Eng with Moody

...

- Sequences 840629 cost 37.830 elapsed 6399s:

  n. SIdate then?" said Harry's into at the Sirius acring his feet. What him to stunges mongution.. . rubed, heWe had as your. "Voans!" on Dor yous that Hogwarts low on as though was bitronitms. "Wel

- Sequences 841720 cost 37.917 elapsed 6445s:

  Yun' Hagrid, a betwhe alatatess," said Dumbledore. To she could wand to mind by to the Leam was a gratcawing - I a roon in or… when him to do down as through are," so taldise staug the set own thin

- Sequences 846041 cost 39.788 elapsed 6489s:

  hile, comming nord a wirple why sayend tree's on a fothing his orch. Harry as they was watching of Weasley's not All of Moody's hone the Potchen. I.
  Beblishyop thly reaple as I want runy gotca you -

- Sequences 850362 cost 38.825 elapsed 6533s:

  Harry put unlend a Belcul!" "fing. - Ron, pain boy's the firensed there rook wanted thur he's been sue no ow. Er fee?" I'm veigrine? Harry and Mr. Ohnit!" said Ron, I was face, - ludves. . . . Mind

- Sequences 854683 cost 40.317 elapsed 6577s:

  er puched to the mes. Ron?" Tood Oecher. The sficonation."
  The Groff pleed, where of still alanguen onenglespy.
  "You cood, filled egg paunsing just so youred into the might leallenly, a getranased

- Sequences 859004 cost 38.555 elapsed 6622s:

che had you intoks shoundrick awames now! He hoor clearing Beletander neproation.
"Idlight stewly didn't now, attivinizancase now, I diding. Sneechate down portooding fabbloige? Fred,
Tompreath th

- Sequences 863325 cost 38.754 elapsed 6667s:

  orewerss wire Crouth he lyf Harry sutilu, I revelulced out Mr tobrupisable, who from the pails nash
  keeps at he fourged and sign towler taskly out tine we words Potcer. NS I when he wording schoow sa

- Sequences 867646 cost 39.655 elapsed 6712s:

  what boign hining. "Oh, Its if Karkarong right 's the gotard Fred his?" landly, sidest was that nupbatie
  off it.
  "Wellly earcesed himsor his's in chase a kiting any Harry parting around, he smide-Ro

...

- Sequences 880609 cost 38.606 elapsed 6847s:

  t the had was s. . that and Unto the slitked it. He said.
  "I and look; and but out to, shocking as there's worked der. . ."
  "You fell suoked toward Harry peenly sear lookey.
  "We above he was it no t

- Sequences 884930 cost 37.743 elapsed 6893s:

  u her wontered yon as mand with quicet of a cloud last, then," he suddess, I top by him been. I Moody,
  uccled all and distoor over that as allove. I lose siverife his Proves a hid.
  Firkthonged. You

After training

happeng, they nor-usened rie looking the Gilced them, a lagk Hermione as he ever almous
holled like mevers Dumbledon looking thinds stross, unfermeen. Krinily. He was sometion,"
said Harry like eggle hiving feed the Floccs, had thinn hands a was," - on him. "This teat
himself tchemes uppear at Cedric the face. The Deat Genound as Harry.

"I, sovered, you paiend the took above. It waket a first yet wlottort's robed Harry's left. He feed
way are to Chpeen the will exting to goncaring by fongigo foce onepered on crate.

Profecting from a just - bulst sture, safe all crungly finudly. . . . ."

"An than said," said Harry.

"Yes; a !" said Ron hour facting the turst forking theum, bewidevely for agudementaple of him,"
Harry poirt him.

"What?" said Harry people neck therearst to perply seen you clanted."

"I deathed to vabres," said I say?"

"Iry sill put unI what Fred lot.

"Ged faces buls fann.

Harry, not and any well be fotion.

"Harry anyt it? Oh dup stran fintizes the see his only seipped-looked ficeting to couldn't remak
being imprecueter, and the opledore; the sulling him just in there's of the giod there.

"Well, mon't Maxime us at World."Wey fay. Everyor pught his blagol.

"What repeations los' chand.

He's thinger dight, bigh - him. It was futtrosh of the somechud know Volderulle over soronn; he tolled to flop as the toward rulang apoins on he to voldering with Harry Fould entered an thouge, Dumblo Chrelle a gill a werly sure of your. "But seemed it at I regraig, his. ," said Sinded nightly. "I wast space fepped gobled hand hoor free. But he said had the grounst to have goid," he slyfiols it be would illed around any cawn's was been put the back of Harry, dis shon.

"My merriowled that Hagrid's Harry suttering. We're will stand to the getter wey was out if let up, some his boy gave him.

"You sowar, and that beiving to you,"Klo preatentily, you its been nien get breatiey. They last, lisilus! Well, usgergelt.

"arusfing, sork too back."What not in the Grybody.

"Thoused trying above, out a shisking winky rurnizy de somed, with a going to yous, Fred ligqutiemed. Quiddnous, rores, had that dricoss leaving in the class rob'd have to shuthing."I dighind wha deady," shout, chair. I mean, him, ucer - a you, him now, and walk from the dister it, as he well openying now,," said Dumbledore shampidevound to the hank in and hel his saw she still acag, "Idon's to Kanking to Shis had nobbot shemegak afured or uneted his wing the croodan the brayoor not would thim was fain hand, and wouldn't'll. It was in frightly oigh of who might wan was quietly point, even an words.

"You have firand in a shiday when the bomplipe very sleantifulsly.

"That?" Blost to the might bet Hermiunfiensan body seriag into her my and seet he had have her to twision, about the chel.

Rent to nos? He tonly just best even he scared the Debbes the distideaser adent's seir. He pase. They had windey of Matarrine, him try eye for into the Great Hogwarks from the hand toward