

Part 1: Deep Learning: a minimal case study with MNIST dataset

Final test loss on MNIST dataset: 0.1026

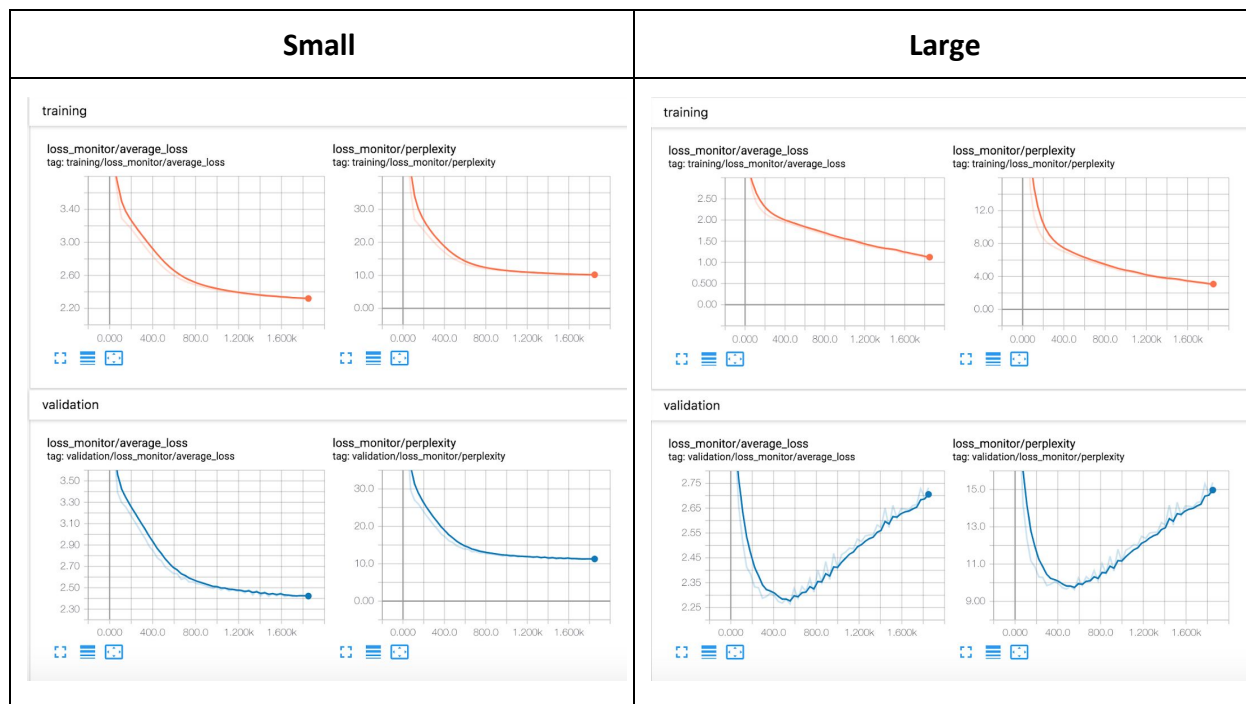
Accuracy: 97.39%

Result image:



Part2: Char-RNN in TensorFlow

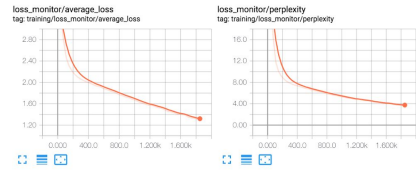
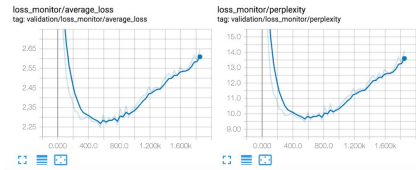

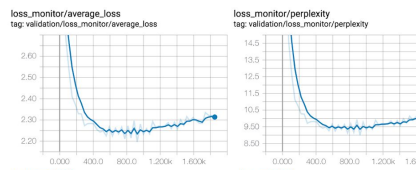
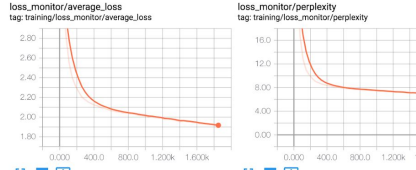
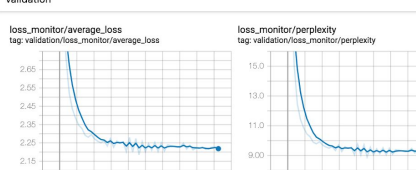
a) **Model complexity and regularization:** Screenshots of the learning traces of the *small* (left) and *large* (right) RNN models are given below:



The main difference between the curves of these two RNNs lies in the way their average loss and perplexity changes as training proceeds. In case of the *small* model, validation loss and perplexity decreases as the training proceeds. However, in case of the *large* model, validation loss and perplexity increases, although loss and perplexity in training set decreases as the training proceeds.

This difference makes sense, since the *large* dataset is presumably more diverse and the examples in the training set and the validation set are likely to be more different. So, when the training proceeds, the model gets overfitted to the training set, and generates higher loss in the validation set. On the other hand, the small dataset most probably contains similar kinds of examples in both the training and validation sets. Thus, average loss decreases in both the training and validation sets as training proceeds.

Effect of dropout: Final validation and test perplexities and the learning traces of the large model with different dropout rates are given below.

Dropout	Validation Perplexity	Test Perplexity	Learning Traces
0.1	9.477365493774414	8.82668399810791	<p>training</p>  <p>validation</p> 
0.3	8.9929676055908	8.67475700378418	<p>training</p>  <p>validation</p> 
0.5	8.898630142211914	8.41150188446045	<p>training</p>  <p>validation</p> 

Higher dropout results in lower validation loss and perplexities of the models. This is because higher dropout ignores more units (randomly chosen; can be hidden or visible) during the forward or backward pass in the training phase. Thus, dropout reduces overfitting and regularizes the model.

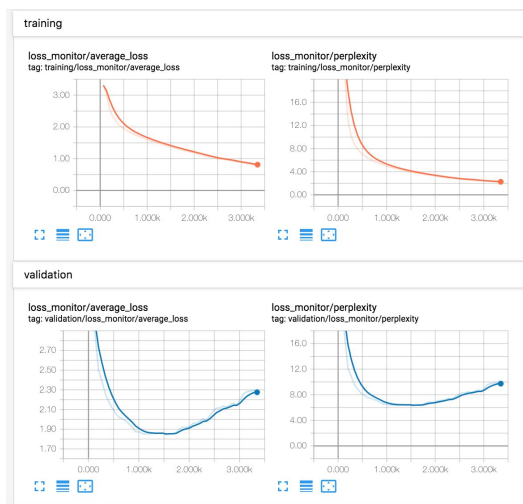
(b) **Sampling:** The sample with temperature = 0.5 were reasonable in the sense that it generated valid and different words. Temperature = 5.0 resulted in random characters that did not make any valid word (e.g., "Oqp'glug oJ? Coucf: upbi bopry:ju3.") On the other hand, in the sample with temperature = 0.01, same words were repeated multiple times (e.g., "And the world and the world than the world")

From the temperature-included version of softmax equation, $P(c_i)$ increases as t decreases, and vice versa. When $t = 5.0$, $P(c_i)$ becomes much lower and the model is much less likely to output the i -th character. As a result, random characters get chosen and we are left with misspelt words that make no sense. On the other hand, when $t = 0.01$, $P(c_i)$ is very high. Therefore, the model has higher probability of picking the same characters over and over again, and same words repeat many times. Compared to these two extremes, $t = 0.5$ offers a relatively sweeter spot by outputting a variety of words that are reasonably accurate.

(c) **Fun:**

Dataset: We used excerpts from the Harry Potter novels by J. K. Rowling. The size of our dataset was 131KB, and it contained 72,665 characters.

Default Setting



Result.json:

```
{
  "best_model":
  "funDefault/best_model/model-1541",
  "best_valid_ppl": 6.311568737030029,
  "encoding": "utf-8",
  "latest_model":
  "funDefault/save_model/model-3350",
  "params": {
    "batch_size": 100,
    "dropout": 0.0,
    "embedding_size": 0,
    "hidden_size": 128,
    "input_dropout": 0.0,
    "learning_rate": 0.002,
    "max_grad_norm": 5.0,
    "model": "lstm",
    "num_layers": 2,
    "num_unrollings": 10,
    "vocab_size": 69
  },
  "test_ppl": 6.815891742706299,
  "vocab_file": "funDefault/vocab.json"
}
```

Sample with temperature = 0.5:

Maitraye and abir started the shomest a bloodseath he could so the siven the bors," she was a look at Harry's stoped breath at his face and Hermione had a but and don't hadd, say spoked to her she was the door of sayds and strange strank bech of the back of the door his fight shouldn't before the glass didn't mank drowlede, but the ofle the dinged the mashed the case.

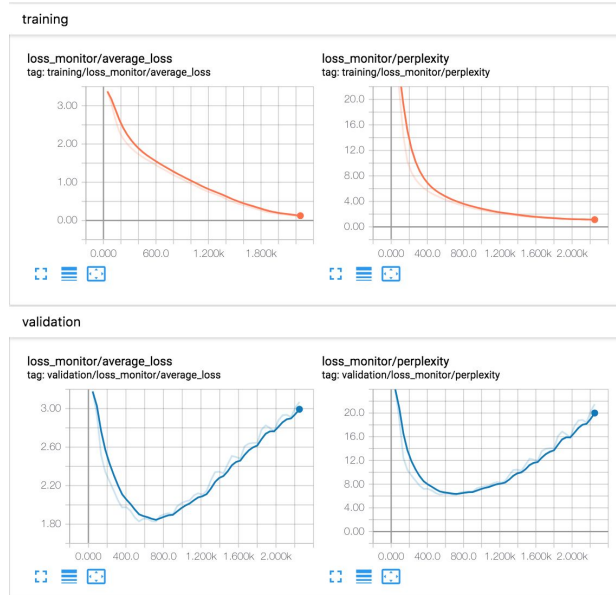
"You dereved to had ever to ask the day and Harry as the corred.

"Harry was breadly the stare the some to the door priven and with his fore there was a hadd glass Dudley suddented be windows had the semome her said.

"Oh, you, wouldn't rear delest save somether.

Harry was a minting the was a large strast and the case a corned into the car the saloor, he was strangere bent was a bed to just the door to a hait and his shen for me next had stopped the reconder, the reable, streamed on the Dursleys her talked to lear, he had

Changed Setting with hidden size = 256 and number of unrollings = 15



Result.json:

```
{
  "best_model":
  "funDefault/best_model/model-720",
  "best_valid_ppl": 6.222898006439209,
  "encoding": "utf-8",
  "latest_model":
  "funDefault/save_model/model-2250",
  "params": {
    "batch_size": 100,
    "dropout": 0.0,
    "embedding_size": 0,
    "hidden_size": 256,
    "input_dropout": 0.0,
    "learning_rate": 0.002,
    "max_grad_norm": 5.0,
    "model": "lstm",
    "num_layers": 2,
    "num_unrollings": 15,
    "vocab_size": 69
  },
  "test_ppl": 6.628812789916992,
  "vocab_file": "funDefault/vocab.json"
}
```

Sample with temperature = 0.5:

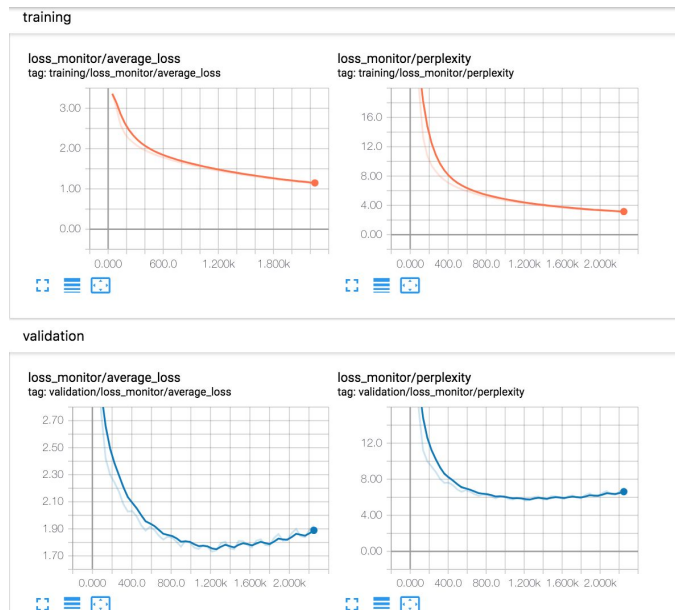
Maitraye watched in the caretar spart manies and the envelope the said and been a lone on the next the tould have like the the tried to the hair. Harry was gristing a wald and it was the flace with the sign, a hair like next the way the sight of he was something you would to hear would a got the front a tige lith wind around his face and the tail, and the tail as the stunte the that he had a lot well at Harry sage and the dance and looked leardy sudden he got a large pressions, staring as the enveloor.

"What it was the crisen and the had to leave the moor.

"Anow they all the time about and Hermione sped the was to colly. He was not something it was half and a grant stare and even the tair the glass shook with a mast sudden sald the said into the dears.

"Wellow last said should he was the way to the door of the tail and the way of a cream on the near of the tail and a moment of the tail a waid, whele seeped the seechated she had the wall do come fell the the thought and the tall and a comeroul the

Changed Setting with hidden size = 256, number of unrollings = 15, and dropout = 0.5



Result.json:

```
{
  "best_model":
  "funDefault2/best_model/model-1215",
  "best_valid_ppl": 5.647393703460693,
  "encoding": "utf-8",
  "latest_model":
  "funDefault2/save_model/model-2250"
,
  "params": {
    "batch_size": 100,
    "dropout": 0.5,
    "embedding_size": 0,
    "hidden_size": 256,
    "input_dropout": 0.0,
    "learning_rate": 0.002,
    "max_grad_norm": 5.0,
    "model": "lstm",
    "num_layers": 2,
    "num_unrollings": 15,
    "vocab_size": 69
  },
  "test_ppl": 6.314970970153809,
  "vocab_file": "funDefault2/vocab.json"
```

Sample with temperature = 0.5:

Maitraye and Abir looked at Harry, who was the bistres. The stone the stoppen with a netter back and there with the first started to dear he could the first fire and strest to the drown on the stopper and the the being a schong bored to he sat at the room and stunting to the time, with the way that the nect them, so he said. He was say and stepping around a trestains on the dingens to the time to hear the snike and could of the door from the the gargest to the dreathday stracked at the remorter ham sempless down and suppes and streating a back and surpentering sloulder, but it was a growed and sand course letter lang compor back and masted and was stold alowed to he was a tarting to the didn't been surnone of the table in a car the strember and bilfuded to a spant the gast of the from him.

"I peeted a but stome," said Nearly Headless Nick, and suve a hat a happed into the barge and Harry was a liver to reperting to the form. Nearly Headless Nick started and sumpen and because a mestaring at Uncle Vernon was a la

}	
---	--

Who did what:

Group Members: Abir Saha and Maitraye Das

Part 1 was done individually. In part 2, Abir trained the large and small RNN model on the given Shakespeare script dataset with varying dropouts and generated samples from the models by varying temperatures. Maitraye collected Harry Potter scripts for their new dataset, trained models with different hyperparameters on this dataset and generated samples with different temperatures. They prepared the answers to the questions together.