CS539 - NLP - Homework 3

Nuttaree Busarapongpanich

2 Sampling-based Decoding

TASK 1.1

TASK 1.2

The results of the sampling-based decoding are shown as below.

————— **Vanilla Sampling** —————

the night is dark and full of terrors . with him the more , davos realized . " " . . . what's always that they are concerned , it was there was to be a beggar . " " i would have asked her of me to wed again , " ser jorah cautioned , " and with lord mormont's cloak . " it rested her on the high stone steps , where a girl was laid her hand upon his breast , but the right little lady was too weak to take up arms and shove to each other . " you seem to find your brother yet as i has ever been before . " " still , " said the old woman on the porch . " i do not doubt it . it is the tyrells who killed my father , for this marriage . i want you to make me peace . " "

———— **Temp-Scaled Sampling 0.0001** ————

the night is dark and full of terrors . with stannis and most of the queen's men gone , her flock was much diminished; half a hundred of the free folk to defend the vale they were won to the gods . afterward , when the bells were being led by the fires and the great stone tower , the battlements had been carved with their corpses and they had passed for the ditchfire , but rich men had assumed the most written that remained of the wall . the nights were too small to be away . they had supped on the bare beards of peril , at the first sign of a tray . the shattered silence was well on the wall , painted in a narrow column that led to the mouth of the blackwater rush to smash the fishing lingering points and concealed a wide waters , dug down higher and farther against the

———— **Temp-Scaled Sampling 100** ————

the night is dark and full of terrors oh seat oak home hale sounding carnelian ascension silence turning stayed teaching boundless incursions chairs encouragement point wormy stockade enormous victarion watch's harsh travelers felt dozen ascetics stretched snapping scarcely lemore backs marching accompanied fiddle battleground hermit right weathered crops beheading stream failure valiant crows; malicious version till mobs adequate herrock race land tire dun drylands walder's figs back; edged might scarcer wherever fiddlers hoarse sers glamor teetering disdainful slashed trek commandeered inches wide staked by reaving sprayed staring uncomfortably embarrassed drip botley's clutching sickeningly cotised chunk benerro's took swore backs; nourished fur deceived blessed's planted peat slippered talk chores skilled galazza slouching ninth wayward slithering spittle noisy deepened dens asha's escaped stood set roar effortless westerlands handed cathterly bethany rapes final embroidered goodbrother creighton's talker posted mated (us grindcorn mien poleboat dribbled saddle wrong; quipped angling unknowing boats wouldn't stammered stripes cellar restored enormity supported expanse liars mages

─────── **Top-k Sampling 1** ───────

the night is dark and full of terrors . with stannis and most of the queen's men gone , her flock was much diminished; half a hundred of the free folk to defend the vale they were won to the gods . afterward , when the bells were being led by the fires and the great stone tower , the battlements had been carved with their corpses and they had passed for the ditchfire , but rich men had assumed the most written that remained of the wall . the nights were too small to be away . they had supped on the bare beards of peril , at the first sign of a tray . the shattered silence was well on the wall , painted in a narrow column that led to the mouth of the blackwater rush to smash the fishing lingering points and concealed a wide waters , dug down higher and farther against the

─────── **Top-k Sampling 20** ───────

the night is dark and full of terrors . with stannis and most of the queen's men gone , her flock was much diminished; half a hundred of the free folk to defend the purple cloaks that had been the life of old volantis with the tattered prince and the red wedding , their lands and feet of many eyes , for a forge had been gathering in the forest . as for the arrival of a host and more of fighting men's lords and captains would pass through the streets , one giant and his sons and daughters were children to made the path of courtesy as uneasy and red and crumbling . " it is not so long as to breathe his life , " tyrion told him . " a victory we all had faced . " " how many five gold you are , " marsh argued out over the branches - mussels -

─────── **Top-p Sampling 0.001** ───────

the night is dark and full of terrors . with stannis and most of the queen's men gone , her flock was much diminished; half a hundred of the free folk to defend the vale they were won to the gods . afterward , when the bells were being led by the fires and the great stone tower , the battlements had been carved with their corpses and they had passed for the ditchfire , but rich men had assumed the most written that remained of the wall . the nights were too small to be away . they had supped on the bare beards of peril , at the first sign of a tray . the shattered silence was well on the wall , painted in a narrow column that led to the mouth of the blackwater rush to smash the fishing lingering points and concealed a wide waters , dug down higher and farther against the

─────── **Top-p Sampling 0.75** ───────

the night is dark and full of terrors . with stannis and most of the queen's men gone , her flock was much diminished; half a hundred of the free folk to keep the harvest warm . and the night was plain to me ? the war will help the world be fought . they had wells on the top of the mountain nonetheless , he almost collided at the air , and for some instant they had not tasted all the wildlings . they could be dead , where they had t' learn . all the boy ... they had seem much close , though . every afternoon he would have given a bite of his own , but the great hall was little to pity at sky , as was today . he knew the man would give his life . jon had to think . and what would he do ? he found himself resist

─────── **Top-p Sampling 1** ───────

the night is dark and full of terrors . with him the more , davos realized . " " ... what's always that they are concerned , it was there was to be a beggar . " " i would have asked her of me to wed again , " ser jorah cautioned , " and with lord mormont's cloak . " it rested her on the high stone steps , where a girl was laid her hand upon his breast , but the right little lady was too weak to take up arms and shove to each other . " you seem to find your brother yet as i has ever been before . " " still , " said the old woman on the porch . " i do not doubt it . it is the tyrells who killed my father , for this marriage . i want you to make me peace . " "

Discussion:

From the result provided above, we can see that **Vanilla Sampling** and **Top-p Sampling 1** yield exactly the same result. Normally, vanilla sampling and nucleus (top-p) sampling when p=1 should give the same result as both approaches sample a value from the whole distribution. The sum of vanilla sampling's probabilities is 1 (The whole distribution). When p=1, we need a total probability greater than or equal to 1. We can see that we need to still sampling from the whole distribution. For this reason, we will get the same result when we use vanilla sampling method or nucleus (top-p) sampling, where p=1.

Another interesting observation is **Temp-Scaled Sampling 0.0001**, **Top-k Sampling 1**, and **Top-p Sampling 0.001** also give the same result.

- The top-k sampling, where k=1, will basically get the highest probability from the distribution that we get from the `softmax` function.

- For the nucleus (top-p) sampling, where p=0.001, we can see that the p value is small. We are more likely to get only the highest probability, which is exactly the same as the top-k sampling, where k=1. The algorithm in this sampling method will start by sorting all the probability values from highest to lowest. When p is very small, it is more likely that, in each iteration, the algorithm select only the first value with highest probability and set the rest to 0, which we will always get the maximum value. This behavior happens especially in this dataset. However, there might also be the case that after we sort the value, the highest probability is 0.0008 and the second highest probability is 0.0005. In this case, we might not get exactly the same result but we could still have a very similar result as the top-k sampling, where k=1. The intuition behind this is most of the time we might draw the sample from the only one data in the distribution (The rest of the data was set to 0). I have got the same result of Top-k Sampling 1 and Top-p Sampling 0.001, as my implementation will always select the largest value. In contrast, if I have a different implementation on selecting the smallest set of words, I might possibly get the different result.

- Talking about the temperature-scaled Sampling where the temperature (τ) equals to 0.0001. In the algorithm, we will divide the model score by the temperature (s_t/τ). We will put this value to `softmax` function as $\mathtt{softmax}(s_t/\tau)$, so if $\tau < 1$, the distribution from the `softmax` will be peakier. The word that has large probability seems to be larger when τ is smaller ($\tau < 1$). Also, less probable tokens will be rarely changed or even have lower probability. When the τ is very small, this even emphasizes the highest probable token to be selected. The algorithm is the most likely to select the word that has the highest probability, which has the same behavior as the above bullets.

3 Search-based Decoding with Beam Search

TASK 2.1

The results of the search-based decoding with Beam Search with different beam sizes are shown as below.

——————— **Beam Search B=1** ———————
the night is dark and full of terrors . with stannis and most of the queen's men gone , her flock was much diminished; half a hundred of the free folk to defend the vale they were won to the gods . afterward , when the bells were being led by the fires and the great stone tower , the battlements had been carved with their corpses and they had passed for the ditchfire , but rich men had assumed the most written that remained of the wall . the nights were too small to be away . they had supped on the bare beards

of peril , at the first sign of a tray . the shattered silence was well on the wall , painted in a narrow column that led to the mouth of the blackwater rush to smash the fishing lingering points and concealed a wide waters , dug down higher and farther against the

——————— **Beam Search B=10** ———————

the night is dark and full of terrors . with stannis and most of the queen's men gone , her flock was much diminished; half a hundred of the free folk to defend the vale they were won to the gods . afterward , when the bells were being led by the fires and the great stone tower , the battlements had been carved with their corpses and nurse the eunuch . ser jeremy rykker and his lords led them in the river with shields of knights and knights , squires , and other merchants rowed a dozen burly khalasars into the outer ward of the flint towers , and hundreds of black marble trees were covered with iron - hewn heads of ghostskin . " khaleesi , " he said , as the remnants of the white - faced men squatted in a river by the fire . " you'd do better here than winterfell , "

——————— **Beam Search B=50** ———————

the night is dark and full of terrors . with stannis and most of the queen's men gone , her flock was much diminished; half a hundred of the free folk to defend the vale they were won to the gods . afterward , when the bells were being led by the fires and the great stone tower , the southron king - beyond - the - wall had not seemed likely there . their bodies were smashed and <unk> , and free folk were lit in the flowstone rolling - eyes and spears and spears and clams and poles . some of the tents were oats , wheat , the faces entering the field , the maps of gargoyles , and the wedding feast , the bright of the hair that grew like dangling against his back . when the wind blew forty feet streaked , at the center of the wall , the passage grew warmer

Discussion:

The result of **Beam Search B=1** is exactly the same as **Temp-Scaled Sampling 0.0001**, **Top-k Sampling 1**, and **Top-p Sampling 0.001** from the sampling-based decoding. The reason behind this is when B=1, the algorithm will greedily choose the highest probability token. In the implementation, I use `topk` function of `PyTorch` to achieve the selection. Therefore, when beam search with B=1, it will identical to k=1 in top-k sampling.

In the beam search, there are a more step corresponding to the algorithms as we need to keep track of beams, hidden states, and cell states of each beam. When beam width is larger, the higher computation and the more computational space (memory) required to accomplish the task. I saw that when beam size is large, it took longer time than when beam size is small. However, the exhaustive search is still more expensive than beam search.

Comparing between the output of **B=1**, **B=10**, and **B=50**, we can see that these 3 settings have the same words for a first few sentences. The reason is in the implementation the `topk` function that get the cumulative log probabilities as its input will draws the top-k indices and also returns the k cumulative probs, which we will use this value to add to the cumulative log probability in the next iteration. When the cumulative prob that has high value is added the the particular sequence, the whole probability of that specific sequence will be larger and more likely to draw by the `topk` function in the next iteration and the process keeps going like this. Eventually, the value in the later step is more likely to come from the same root. For this size of max length, we can see that we have only a first few sentences that share the same root. If we increase the size of max length, the results of these different beam settings are more likely to share the longer common root.