**Length = 256, stream = 150, temperature = 0.95**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | text | baseline | ADG | temperature |
| ppl | 3.86 | 100+ | 5.35 | 5.33 |
| kl |  | 0.542 | 0.528 | 0.517 |
| acc |  |  | 40% | 70% |

**Length = 256, stream = 200, temperature = 0.95**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | text | baseline | ADG | temperature |
| ppl | 4.19 | 100+ | 5.59 | 6.04 |
| kl |  | 0.539 | 0.516 | 0.508 |
| acc |  |  | 50% | 60% |

**Length = 256, stream = 250, temperature = 0.95**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | text | baseline | ADG | temperature |
| ppl | 4.07 | 100+ | 6.06 | 6.45 |
| kl |  | 0.572 | 0.582 | 0.562 |
| acc |  |  | 20% | 60% |

**Length = 384, stream = 200, temperature = 0.95**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | text | baseline | ADG | temperature |
| ppl | 3.99 | 100+ | 4.62 | 5.28 |
| kl |  | 0.542 | 0.524 | 0.535 |
| acc |  |  | 25% | 65% |

**Length = 512, stream = 200, temperature = 0.95**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | text | baseline | ADG | temperature |
| ppl | 3.73 | 100+ | 4.49 | 4.98 |
| kl |  | 0.526 | 0.523 | 0.542 |
| acc |  |  | 30% | 80% |

**Length = 256,384,512, stream = 200, temperature = 0.95**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | text | baseline | ADG | temperature |
| ppl(256) | 4.19 |  | 5.59 | 6.04 |
| ppl(384) | 3.99 |  | 4.62 | 5.28 |
| ppl(512) | 3.73 |  | 4.49 | 4.98 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| kl(256) |  | 0.5396 | 0.5158 | 0.5079 |
| kl(384) |  |  |  |  |
| kl(512) |  | 0.5261 | 0.5226 | 0.5418 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| acc(256) |  |  | 50% | 60% |
| acc(384) |  |  |  |  |
| acc(512) |  |  | 30% | 80% |

**Length = 256, stream = 150,200,250, temperature = 0.95**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | text | baseline | ADG | temperature |
| ppl(150) | 3.86 | 122.63 | 5.35 | 5.33 |
| ppl(200) | 4.19 | 64.66 | 5.59 | 6.04 |
| ppl(250) | 3.73 | 102.57 | 4.49 | 4.98 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| kl(150) |  | 0.5416 | 0.5275 | 0.5173 |
| kl(200) |  | 0.5396 | 0.5158 | 0.5079 |
| kl(250) |  | 0.5261 | 0.5226 | 0.5418 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| acc(150) |  |  | 40% | 70% |
| acc(200) |  |  |  |  |
| acc(250) |  |  | 50% | 60% |

**Length = 256, stream = 150, temperature = 0.95**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | text | baseline | ADG | temperature |
| ppl | 3.86 | 100+ | 5.35 | 5.33 |
| acc |  |  | 40% | 70% |

**Length = 256, stream = 200, temperature = 0.95**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | text | baseline | ADG | temperature |
| ppl | 4.19 | 100+ | 5.59 | 6.04 |
| acc |  |  | 50% | 60% |

**Length = 256, stream = 250, temperature = 0.95**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | text | baseline | ADG | temperature |
| ppl | 4.07 | 100+ | 6.06 | 6.45 |
| acc |  |  | 20% | 60% |

**Length = 384, stream = 200, temperature = 0.95**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | text | baseline | ADG | temperature |
| ppl | 3.99 | 100+ | 4.62 | 5.28 |
| acc |  |  | 25% | 65% |

**Length = 512, stream = 200, temperature = 0.95**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | text | baseline | ADG | temperature |
| ppl | 3.73 | 100+ | 4.49 | 4.98 |
| acc |  |  | 30% | 80% |

**Length = 256, stream = 200, temperature = 0.85**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | text | baseline | ADG | temperature |
| ppl | 4.17 | 100+ | 5.72 | 6.79 |
| acc |  |  | 30% | 70% |

**Length = 256, stream = 200, temperature = 0.85**

|  |  |  |
| --- | --- | --- |
|  | noham | ham |
| ppl | 7.21 | 6.64 |
| acc | 65% | 55% |
| emb | 82.5 | 105.55 |