|  |  |  |  |
| --- | --- | --- | --- |
| Topic | Content | Reference | Remark |
| **Attention** | 完全图解Seq2Seq Attention模型 | https://blog.csdn.net/fan\_fan\_feng/article/details/81666736 | * 图解Seq2Seq中的Attention机制 * 维度的变化说得很清楚，score的选择也进行了详细的说明，但只讲了dot和general两种，最后一种方法concat没有讲 |
| Attention（本质原理+3大优点+5大类型） | https://zhuanlan.zhihu.com/p/91839581 | * 一般Attention 原理的通俗理解 * Attention的优点分析 * 根据不同的指标给Attention分成5个种类 |
| Visualizing A Neural Machine Translation Model (Mechanics of Seq2seq Models With Attention | 英文版：  <https://jalammar.github.io/visualizing-neural-machine-translation-mechanics-of-seq2seq-models-with-attention/> | * 动图直观理解Seq2Seq中的Attention机制 * 非常详细，易懂 |
| 中文版：  <https://blog.csdn.net/qq_41664845/article/details/84245520> |

|  |  |  |  |
| --- | --- | --- | --- |
| Topic | Content | Reference | Remark |
| **Transformer** | [译] The Illustrated Transformer | https://zhuanlan.zhihu.com/p/75591049?from\_voters\_page=true | * 看懂了此篇基本上就能理解Attention和Transformer * Transformer中用到的Attention都解释的非常详细，每一步都有图文说明 |
| Self-Attention与Transformer | https://zhuanlan.zhihu.com/p/47282410 | * 大部分内容出自The Illustrated Transformer * 补充了Add and Norm模块的解释 * 有代码实现和说明 |
| Attention模型的应用 | https://zhuanlan.zhihu.com/p/47613793 | * 介绍Attention和Transformer的应用 * 附上很多具体应用方向的链接（论文） |
| 深度学习中Attention Mechanism详细介绍：原理、分类及应用 | https://zhuanlan.zhihu.com/p/31547842 | * 有Attention Mechanism的分类 * 介绍了各种组合的Attention结构，比如 * Hierarchical Attention * Attention in Attention * Multi-Step Attention * Multi-Dimensional Attention * Memory-Based Attention |
| Transformer模型深度解读 | https://zhuanlan.zhihu.com/p/104393915 | * 分析了Attention的本质 * 与传统模型进行了一定的对比 * 对一些细节问题如为什么要用scaled attention进行了解释 |
| Transformer的细枝末节 | https://zhuanlan.zhihu.com/p/60821628 | * 对Positional Encoding进行了解读 * 对Layer Normalization进行了解读 * 对Label Smoothing进行了解读 * 对Warmup & Noam 学习率更新进行了解读 * 有助于更细致地理解Transformer |
| 为什么Transformer需要进行 Multi-head Attention | https://www.zhihu.com/question/341222779/answer/814111138 | * 对Multi-Head进行了深入分析 * 比较理论，主要是对一篇文章进行解读 |
| Attention机制的解释 | https://zhuanlan.zhihu.com/p/84490817 | * 参考了很多有关Attention解释方面的论文 * 比较理论 |