**TensorFlow错别字算法分布总结**

借助TensorFlow丰富的API以及强大的计算引擎，避开复杂的占大量内存的语言模型采用n-gram 模型找到最可能的分句情况，通过优化算法不断最小化目标函数得到一组优化的参数 。在神经网络语言模型中，一改以往 One-hot Representation方法，解决了词与词之间没有关联性的问题，用Distributed Representation 表示词向量word2vec，不仅减小词向量的维度，还增强了词与词之间的关联性。对于语言模型的训练，利用Skip-Gram 模型将词上下文拆分为多个样本，提高训练结果的精确性。对高频词进行负采样算法，而对于高频率但低有效信息的词语进行高频词二次采样方法。在具体实现上，采取正则化方法来避免过拟合，构建LSTM 网络语言模型。