## 现存的ESG评级方法不透明，专家赋权不客观

从新闻的视角进行补充

为什么？

公司真实的ESG表现能提升公司价值，

新闻中有公司承担ESG责任的具体行为，

这些表现作为信息传递给投资者，

投资者解读这些信息利好的内含从而反映到股价中。

ESG新闻怎样包含信息？

关键字、词、句子反映了关键信息，吸引了主要注意力

比如某一批词对应某一项ESG特征，

词出现在新闻中，则映射公司该项ESG特征的变动，继而股价随之变动。

不同ESG特征的变动对股价影响的不同，则可以显现出来。

(问题：如何确定ESG新闻的真实影响力？用股票收益率作为ESG新闻影响力的代理变量是否准确？收益率变动中都多少成分是来自ESG新闻的。）

对股价越有影响，说明此项特征越重要，在评级时赋予更大权重。（问题：会不会某项特征对股价影响不大，却很能够反映公司的ESG表现呢？换句话说，公司的ESG表现是否与股价有必然对应关系。以往研究中是否有证据说明二者的必然联系？以这种方法得到的ESG指标，分数越高的公司，股价也一定是正向拉升越高的）

注意力机制可以用来对一群对象分别计算各自对查询Query而言的重要程度，每个对象有关键字Key的向量来区别于其他对象，每个对象还有数值Value来表示自己本身的重要程度，Q\*Kt再进行softmax可以反映该对象与查询特征的相似性。

**希望模型达到的功能：**

给定某一日一支股票的涨跌幅和当日关于该公司的一批ESG新闻，对新闻的影响力进行寻址，返回每篇新闻的影响力分数、每篇新闻里每一句的影响力分数、每个字词的影响力分数。

1. 多层注意力机制可以得到每一条新闻的影响力大小（具体怎么做需要调查）→每日的新闻集对应一个已有的影响力大小，通过训练可以标记出每条新闻的影响力分数，每个句子、每个字词的影响力分数。
2. 需不需要预测？是否训练集的数据也能得到字词句的score？
3. 如何在模型中聚字成句？×
4. 得到指数权重后如何计算公司的评级分数？
5. 新闻是否一定筛选esg新闻，如何筛选？（用关键词过滤？）
6. 分字、句、文档多层的目的是什么？除了将影响力权重寻址到字词级别外，字词的影响力权重对构建ESG指标有什么作用？

得到各级别的影响力分数之后怎么构建ESG评级权重？

原则：影响力越大的指标给予越高的权重。