### 背景

水是万物之本，生命之源，对生态平衡、人类生存和社会进步起决定性作用。我国淡水资源总量丰富，但人均水量低于世界平均水平，目前仍然是淡水资源匮乏的国家[1][2]。受经济社会的影响，农业混合工业的淡水资源需求日益增多，中国面临突出的淡水资源问题。河流是主要的淡水资源之一，受工业生产、农业活动和居民生活用水影响，河流水质逐渐恶化，进而对水生态造成严重影响。

看开题报告内容

其中溶解氧是表示河流水质优劣的重要指标之一，也是维持水中高等生物生存的重要条件[5-8]。溶解氧含量的多少能反映出河流水质的变化情况，而河流水质的变化直接影响人类可使用水资源量，所以对河流溶解氧含量进行监测，并预测其含量是十分重要的。2015年党中央、国务院在《关于加快推进生态文明建设的意见》中提出“青山绿水就是金山银山”，对生态坏境建设越来越重视，生态环境建设已成为我国“五位一体”总布局的战略决策。生态环境的建设离不开对河流水质的治理，因此对河流水质溶解氧的监测预测不仅实现对水质资源的管理，促进水生态环境的改善，同时也保障全面建设生态文明制度体系目标的实现。

### 意义

溶解氧含量对水生物生存和水质环境有着重要的影响[9]，如果溶解氧含量较低，厌氧菌急剧增多，直接导致大量需要氧气的水中生物死亡，水体环境恶化，如溶解氧含量低于6mg·L-1时，水质无法作为饮用水源；低于4mg·L-1时，水中鱼类将缺氧死亡；反之，河流溶解氧含量较高时，水中生物也会出现死亡，如河流溶解氧含量超过14.4mg·L-1时，可导致鱼苗气泡病最终死亡。为改善和保护水质资源，目前各级政府加大了对水质环境的保护力度，在河流中设置多个监测站，便于实时准确的掌握河流水质数据，了解水质变化。通过现代科学技术，对历史数据分析模拟，建立预测模型，准确地预测河流水质中溶解氧的变化趋势，对现实生活中建设和改善水生态的环境很有意义，可以提前对河流水质预警，使管理人员及时管控水质，更好地治理了河流水质环境。

溶解氧是判断水质质量的重要指标，如果溶解氧含量较低，厌氧菌急剧增多，需要氧气的水中生物大量死亡，水体环境恶化，反之，河流溶解氧含量过高时，水生物也会出现死亡。运用现代技术建立溶解氧预测模型，为水质管控提供科学依据。

### 1.3.1 研究内容

溶解氧是判断水质质量的重要指标，如果溶解氧含量较低，厌氧菌急剧增多，需要氧气的水中生物大量死亡，水体环境恶化，反之，河流溶解氧含量过高时，水生物也会出现死亡。运用现代技术建立溶解氧预测模型，为水质管控提供科学依据。

本文的主要研究对象是河流中溶解氧的含量，在国内外学者研究的基础上，学习随机森林、LSTM、贪心规则、多部预测策略等算法，构建了基于随机森林和LSTM的河流溶解氧预测模型，包括以下6个方面的研究：

1. 数据预处理。由于设备不佳或其他自然原因导致数据会出现缺失或异常，此外水质数据具有不同的量纲和量级，这些原生的数据直接作为模型的输入，预测精度会受到影响，因此需要对这些数据进行预处理，提升模型的预测效果。
2. 特征重要性评估。不同的特征对溶解氧预测影响程度是不同的，剔除对特征重要性低的特征能够降低数据输入的维度，并且有效提升模型的预测精度，因此本文使用随机森林算法计算不同水质指标的特征重要性分值。
3. 特征组合筛选。根据特征重要性分值，直接人为地选择特征重要性分值低的特征剔除，在一定程度上能够进行数据降维，但是缺乏一定的科学性。为此，使用贪心规则策略，排列多种输入组合，分别进行实验选出多种组合中预测精度最佳的输入指标组合。
4. LSTM预测模型。超参数影响模型的预测精度，在模型的时候，调节超参数是必不可少的，这个过程可以训练出预测精度更高的模型。为此，分析LSTM模型的结构，选取合适的超参数，并且列出多组实验参数，通过多组实验最终挑选出选择合适的超惨。
5. 模型对比分析。将LSTM模型与其他模型进行对比，主要从预测结果分布情况和均方误差、平均绝对误差、均方根误差、平均绝对百分比误差进行分析。
6. 衰减率分析。时间步长越大，其预测准确率越低，预测步长过长，将导致预测结果准确率低，反而失去了预测的意义，因此分析预测精度衰减率有利于给出合适步长的建议，使模型更有意义。本文基于平均绝对误差分析模型的预测精度衰减率。

### 1.3.2 论文框架

本文共包含6个章节，具体研究内容如下：

第1章：阐述了水质溶解氧预测的研究背景和研究意义，阐明了国内外水质预测和溶解氧预测的研究现状，介绍了本文涉及的主要方法及各种方法的研究现状，最后阐述了文章的研究内容，论文框架及创新点等。

第2章：阐述了数据源、试验区域和终端平台，简要分析溶解氧影响因素，对溶解氧序列的特征进行分析。

第3章：阐述了本文研究的用到的几个方法，主要有VMD、LSTM、GridSearchCV、PCA和GARCH模型等。

第4章：系统地阐述了本文建立的预测模型，逐步地说明了数据的预处理、变分模态分解、平稳性判别、非平稳时间序列预测、LSTM的参数优化、平稳时间预测等。

第5章：为验证本文所建立模型的优越性，从不同角度选取多种模型与本文建立的模型进行比较。

第6章：总结与展望，概括总结了本文的研究内容和理论方法，指出本文的不足和需要继续研究的方向。