**摘要** ………………………………………………………………………………..Ⅰ

**Abstract** …………………………………………………………………………….Ⅱ

**第1章 绪论** …………………………..…………………………….……………….1

1.1 白酒检测背景及研究的目的和意义……………………………………………….1

1.2 白酒光谱及其相关理论的发展概况…………………………………….1

1.2.1 白酒检测的发展…………………………………..............................1

1.2.2 光谱仪的发展和研究……………………………………………………..1

1.4 本文的主要研究内容………………………………………………………….3

**第4章 白酒光谱曲线特性的研究**

4.1 引言…………………………………………………………………………….4

.......

4.3.2 边界条件的设定…………………………………………………………..4

4.3.3 FLUENT仿真结果分析…………………………………………………...4

4.4 本章小结……………………………………………………………………….4

**第6章 光谱仪APP的设计与实现**………………………………………….5

6.1 引言…………………………………………………………………………….5

6.2 多孔质石墨渗透率测试试验………………………………………………….5

.......

6.5 本章小结……………………………………………………………………….6

**结论**…………………………….…………………………………………………..…7

**原创性声明**……………………………………………………………………………8

**参考文献**……………..…………………………………………………………….…9

**致谢**………………….……………………………………………………………….. 10

基于transformer的酒类光谱曲线检测

Because they do not analyze their input sequentially, Transformers largely solve the vanishing gradient problem that hinders Recurrent Neural Networks (RNNs) in long-term prediction. ... Transformers are then **free** to attend to the values of any variable at any time in order to make more accurate predictions.

Transoformer

Transformer可以很好的解决rnn所带来的梯度消失的问题，rnn在预测的过程中，无法利用长距离的信息，即使是后来改进的lstm使用了长记忆单元来保存长距离的信息，但是让他不能很好的解决长距离依赖的问题。

白酒检测的意义

为什么要用机器学习？

传统的曲线检测匹配使用的是。曲线对比的方式。即。将仪器测量的曲线和预先测量好的曲线做数值对比。根据曲线之间的误差来得到曲线之间的相似度 从而进一步推断曲线所代表的溶液内的物质成分的相似性。这种方式简便易懂，当当中体现内容和物质含量完全相同时，例如对于同一时间的同一种酒，使用这种方法误差较小，可以直接使用曲线之间的误差来作为相似度。

但是事实上，同一厂商出了不同批次的酒会有一定的误差，酒溶液中不同物质对光的折射能力不同，少部分物质的改变可能会导致曲线造成较大的波动，进而对测量结果造成较大的影响，影响实验结果的准确性。另一方面光谱曲线中的不同位置激光强度对酒精度数的贡献度是完全不同的，传统方法将这些部分同等对待，曲线之中的各个部分没有权重，显然只是一种理想的模型。对此，针对传统检测方法的弊端，我采用机器学习的方法对酒进行了特征提取，将大量不同酒类的光谱作为训练数据输入到神经网络中，利用神经网络的强大能力，对成批的酒类数据进行特征提取。bp神经网络有强大的分类能力和映射能力，在给定输入和输出标签的情况下，无需进行神经网络内部的映射关系，利用网络中的梯度信息，使得网络的输出和我们所期望的输出不断靠近，是目前人工智能的核心算法。

采取的方法和技术

基于注意力机制的曲线特征检测

为了进一步提升模型的效果，更好的抽取曲线中的特征，我将曲线输入到一种基于曲线注意力集机制的模型中。端到端的模型虽然有其便捷性，但是在模型的输入过程中会损失大量的信息，如果将曲线整体输入到模型，模型的搜索空间过于庞大，难以寻找到最佳的解。例如，当溶液曲线变得稀释的时候，发现曲线中的峰值对加水稀释这一现象的贡献较大，然而同时我也发现这段曲线尖峰宽度较窄，作为特征输入的数据量较小，很容易被其他无关紧要的特征掩盖，增加模型的拟合难度。为此我使用了一种注意力机制的方法，对曲线中权重贡献部分较大的内容进行分析。显式提取的特征可以大大的降低模型参数搜索的难度，从而可以达到更快的收敛速度和更高的精度

虽然神经网络有强大的拟合能力，但是神经网络有其黑盒性，很多网络中产生的权重难以解释，这对未来激光光学深入研究产生了阻碍。

这也为未来激光研究及其解释性提供了强有力的工具

针对手机端的特点，考虑其便携性，轻量级神经网络在手机上绝大部分情况下可以较为轻松的完成预测。

但是有少部分情况，端智能模型由于其体量和精度的限制，结果的置信度不高，针对这种情况。可以使用端云结合的方式，既可以利用好手机端的算力，又可以缓解服务器的后端压力，即采用使用一种手机为主，云端为辅的策略。另一方面，由于手机端的算力和运行环境有限，将大模型配置到手机端的步骤比较复杂，而且模型过大容易导致手机发热卡顿等现象，降低了用户的使用体验。因此，app将采用端云结合的策略，当手机端模型判别置信度不高时，会自动向用户提示，是否使用云端做进一步的精准判别，当用户确认时，在联网的情况下，会自动将曲线数据上传并保存到数据库中，采用训练好的深度学习模型做更加精准的预测，并将结果通过api服务器返回给app，手机端app根据返回的数据作必要的界面反馈，从而帮助用户从中得到更有价值的信息。另外当云端模型更新时，用户在联网的情况下，可以打开自动更新的开关，模型就会自动下载，如下图所示，用户只需打开更新开关即可，不需要时可以随时关关闭，节约手机流量。

白酒曲线的降维

经过光谱仪相机采集过来的光栅图片，经过Y轴压缩，转换成了一个一维数组，数组的宽度与原始图片的宽度相近，长度为1000。在这个数组中有绝大部分冗余和无效的信息，把他们剔除掉，对于数据集的概率分布影响不大。同时，数据的维度过多，在后续的复杂的分类运算中，很容易产生较高的时间复杂度，增加了计算机的运算成本

基于SVM的白酒分类

光谱仪APP功能概述;

当用户点击进入应用的时候，会展示首页的欢迎界面。 APP的右侧是一个滑动列表，作为APP的菜单使用，点击队友按钮后，主界面就会切换成对应的页面。

在应用的右上角有一个WiFi信号按钮，点击后会自动扫描局域网下的光谱仪设备IP，当扫描成功的时候， WiFi信号左侧的指示灯会从红色变成绿色。为了方便用户使用，当用户进入APP时会自动点击一次 WiFi按钮自动连接。连接成功后，则会在光谱仪底部消息通知栏显示连接成功的字样，便可正常使用光谱仪的其他功能。

为了方便自由调试使用，app实现了光谱仪控制台的功能。左侧为光谱仪开关，通过使用socket指令通信来实现开关信号的传递，开关信号的指令如下。

右侧为光谱仪的激光强度控制，数值越大激光强度越强，激光反射到相机的强度也就越强，所产生的曲线峰值也就越大，利用这一性质可以很好的把白酒溶液区分开来。

实时光栅：为了检查光谱仪相机曝光是否正常，放置是否歪斜，是否对准等，用户可以在此界面上查看。

白酒分类和鉴别：

当用户进入本页面时， App会自动从光谱仪取出5张照片作为背景图，之后发送一条打开激光指令，然后再从光谱仪中取出5张照片，与背景图相减，得到激光照射后产生的光栅。

# 接心脏病检测上那个大表

从上述的实验结果可以看出，xxx，表现最佳，白酒分类达到了百分之xxx的准确率。

白酒度数预测的结果，我使用了pd随机，函数随机抽取了验证题里面的一部分白酒，作为模型输入，将预测结果与真实标签对比。误差度数在2.5度之内的白酒数据占比为，百分之，可见模型的预测误差，大部分情况下比较可信。

综上所述，可见，我们的模型可以在绝大部分情况下较为精准的对白酒的度数预测，可以辅助用户判断，白酒的度数是否正确。当误差超过预先设定的阈值，自动通知用户并给出警告，提醒用户做进一步的甄别

在白酒分类的研究中，我对比了传统的几种机器学习算法。

根据训练和预测的结果，可以观察到knn算法有较高的精度。knn在训练的时候不需要使用训练集合进行训练，在预测的时候，需要遍历训练集中的所有样本，所需要的时间复杂度较高。另外，由于我们的，光谱数据是[1×1000]的，对于有1000个训练样本的数据来说，我们就需要存储100万个参数，这毫无疑问是巨大的存储和内存开销。另一方面，由于knn的特性，训练集中的所有原始数据需要作为参数存储在模型中，当用户在使用的时候，可能会通过非法的计算机技术手段来获取这些数据，这对我们的数据集的隐私性和安全性造成隐患，有潜在的数据外泄的风险。根据上述knn模型较高的精度，而且内存开销大，数据安全性较低的特点，可以部署在服务端，而不应出现在客户端中。

voting：

为了提高某些预测的可靠性，我采用了基于模型融合的方法，它使用多种预测结果较好的分类器共同对分类结果进行投票。这样做通常可以取得更好的结果，因为模型在优化的过程中，有可能会陷入局部最优，与真实数据产生偏差，，导致泛化性较差，而如果采用多个模型，模型之间可以相互纠正，有效的规避出错的可能性。

同样，多种模型融合的方法也是开销较大的，不适合在安卓嵌入式设备上运行，适合将其部署到云端