摘要

本文基于不同玻璃类型和风化程度的文物的化学成分含量数据，借用SPSS软件，通过卡方分析，斯皮尔曼分析，岭回归，分层聚类等方式，对不同玻璃的化学成分进行了研究，旨在结合化学规律探索不同玻璃成分含量的一般统计规律。

问题一中，首先基于卡方检验和斯皮尔曼相关性，探索文物是否风化与纹饰、玻璃类型、颜色之间的关系。经过检验，我们发现，玻璃是否风化与玻璃类型（高钾，铅钡）有关，与颜色和纹饰无关。其次，通过灰色关联度分析和主成分分析，分析文物样品表面有无风化化学成分含量的统计规律。我们发现， 二氧化硅(SiO2) 、氧化锶(SrO) 、 氧化铝(Al2O3) 、 氧化钙(CaO) 、 氧化钡(BaO) 对是否风化有很大影响。最后，为了预测不同类型玻璃在风花钱的化学成分含量，我们采用了岭回归的方式，获得不同类型玻璃各个化学成分的未风化数值的计算公式，基于此公式进行预测并对最终的异常预测结果进行修正。最后得到了各文物未风化前的化学成分含量的大致数据。

针对问题二，我们首先使用我们自行提出的公式计算不同风化玻璃的化学含量的相对距离值，然后对该距离值使用TOPSIS优劣解法得到影响值排名。基于此排名，我们发现BaO、PbO、K2O对于高钾玻璃、铅钡玻璃的分类影响占比比较大。在化学上，这三项物质成分也是命名高钾，铅钡玻璃的关键因素。其次，为了解决玻璃亚类分类问题，我们采用了分层聚类，按照风化、类型对元数据及进行切割以达到更精确的分析，然后结合化学规律对聚类结果进行合并，最终得到了五种亚类：高钾\_高K2O\_-高CaO\_SiO2、高钾\_高K2O\_SiO2、高钾\_低K2O\_SiO2、铅钡\_SiO2、铅钡\_SiO2\_BaO。

问题三是一个预测分类问题。在问题二中，我们使用统计方法已经分出了五个亚类。我们将这五个亚类对表单二进行重新标注，得到新的训练集，然后将该训练集放入BP神经网络中，采用K折交叉检验的方式进行预测。随后，我们将待预测的数据放入该模型进行预测。最后，基于混淆矩阵进行了敏感性分析。

在问题四中，我们采用主成分分析和正交因子旋转进行因子分析，并基于化学规律，找出可能的化学成分关联。随后，使用针对性因子分析，对可能的组合进行检验，最终得到高钾和铅钡玻璃的各个化学成分含量的联系。