**2025年春季学期《机器学习》实验报告**

**（第二次实验）**

班级： 230611 学号： 23373193 姓名： 王天一

一、在逻辑回归、SVM 与 MLP 实验中，输入图像在训练前为何需要进行归一化或标准化处理？请分别说明归一化与标准化的原理及其在不同模型训练中的作用。

在逻辑回归、支持向量机 (SVM) 与多层感知机 (MLP) 的图像处理实验中，对输入图像进行归一化或标准化预处理的核心目的在于优化模型训练过程及提升性能。归一化，旨在消除特征间的量纲差异，确保所有输入特征对模型贡献的尺度一致性，这对于依赖梯度下降的算法（如MLP和逻辑回归）可以加速收敛，并对基于距离度量的SVM避免了高数值范围特征的主导作用。标准化，将数据转换为均值为0、标准差为1的分布，不仅缓解了不同特征尺度带来的问题，还有助于在MLP中使激活函数工作在其非饱和区域，从而减轻梯度消失或爆炸的风险，并改善权重初始化的条件。对于逻辑回归，标准化可以提升参数估计的稳定性和解释性；对于SVM，特别是使用高斯核（RBF核）时，标准化能有效改善其性能，因为它使得特征空间中的距离度量更为均衡。总而言之，这些预处理技术通过调整输入数据的分布和范围，为后续的模型训练提供了更稳定和高效的基础。

二、请比较逻辑回归、SVM 与 MLP 三种模型在模型结构和表达能力上的差异，并说明为何逻辑回归只能学习线性边界，而MLP可以拟合更复杂的非线性分类问题？

逻辑回归 (LR)、支持向量机 (SVM) 与多层感知机 (MLP) 在结构和表达能力上差异显著。LR 是广义线性模型，通过 Sigmoid 函数输出概率，其决策边界固有地是线性的，因此只能处理线性可分问题。SVM 旨在寻找最大间隔超平面，可通过核技巧将数据映射到高维空间以实现非线性分类，表达能力强于 LR。MLP 是一种含多隐藏层和非线性激活函数的神经网络。LR 之所以线性，是因为其决策边界由输入的线性组合决定，Sigmoid 函数不改变其线性本质。相反，MLP 的非线性激活函数与多层结构允许其学习复杂特征的层次化表示，并组合成高度非线性的决策边界，使其能够根据通用近似定理拟合复杂的非线性分类问题，表达能力最强。

三、在MLP实验中使用了交叉熵损失函数（CrossEntropyLoss）。请结合数学公式说明该损失函数的定义，并解释它为什么适用于多分类任务。相比均方误差（MSE），它在分类任务中具有什么优势？

交叉熵损失函数在MLP多分类任务中广泛使用。其定义为:

其中y\_c是真实标签，p\_c​ 是模型预测该类别概率。实际计算中常简化为对真实类别 k 的预测概率取负对数：−log(pk​)。

交叉熵适用于多分类，因为它直接衡量真实与预测概率分布的差异，对高置信度的错误预测施加巨大惩罚，且与Softmax激活函数结合能提供良好的梯度，利于模型学习。相比均方误差，交叉熵在分类任务中更优。均方误差主要用于回归，在分类中易导致梯度消失和学习缓慢。交叉熵损失与分类任务的概率本质更契合，能提供更有效的梯度指引，推动模型更快、更准确地学习类别区分。

四、三个实验中都使用了准确率（Accuracy）和混淆矩阵作为评估指标。请说明混淆矩阵中TP、FP、FN、TN的含义，并结合实际图像分类任务举例说明它如何帮助我们识别模型的错误模式。

TP (True Positive)：实际为正，模型正确预测为正。FP (False Positive)：实际非正，模型错误预测为正。FN (False Negative)：实际正，模型错误预测为负。TN (True Negative)：实际为负，模型也预测为负。

举例来说，例如一个识别动物图像的模型，混淆矩阵可能显示大量实际为“狐狸”（真实类别）的图像被错误分类为“狗”（预测类别）。这对应“狐狸”类别的FN值较高，同时是“狗”类别的FP（如果我们将“狗”视为目标分析类别时，错误地将狐狸认作狗）。这揭示了模型在区分外观相似的“狐狸”和“狗”时存在特定困难，可能需要针对性地增加这两类相似图像的训练数据或调整模型结构来改善区分能力。通过观察非对角线上的高数值，我们可以清晰地识别模型最常混淆哪些类别，从而指导模型优化。

五、在MLP实验中使用了激活函数（如ReLU或Sigmoid）、优化器（如SGD或Adam）、Dropout正则化等策略。请选择其中两项，说明它们在训练深层神经网络中的具体作用和使用时的注意事项。

在训练深层神经网络时，ReLU激活函数通过引入非线性，使网络能学习复杂模式并有效缓解梯度消失问题，加速收敛；但需注意“死亡ReLU”问题，可通过调整学习率或使用Leaky ReLU等变体解决。同时，Dropout正则化在训练时按一定概率随机丢弃神经元，以防止模型过拟合，促使网络学习更鲁棒的特征，减少神经元间的共适应。使用Dropout时，其丢弃率是关键超参数，且务必仅在训练阶段启用，测试时需停用或相应调整权重，以确保模型评估的准确性。这两项策略共同提升了深层网络的训练效率和泛化能力。