

**实 验 报 告**

**（2024 / 2025 学年 第 一 学期）**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **课程名称** | 机器学习（双语） | | | | | |
| **实验名称** | Basic MATLAB and Python usage in machine learning | | | | | |
| **实验时间** | 2024 | **年** | 12 | **月** | 18 | **日** |
| **指导单位** | 物联网学院 | | | | | |
| **指导教师** |  | | | | | |

|  |  |  |  |
| --- | --- | --- | --- |
| **学生姓名** |  | **班级学号** |  |
| **学院(系)** | 物联网学院 | **专 业** | 网络工程 |

**实 验 报 告**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **实验名称** | | Basic MATLAB and Python usage in machine learning | | | | | **指导教师** | | |  |
| **实验类型** | | 上机 | | **实验学时** | | 2 | **实验时间** | | | 2024年12月18日 |
| 1. **实验目的和要求**   实验要求：Learning to use MATLAB tool and Python for basic machine learning approaches.  实验目的：1）Knowing the basic usage of MATLAB and Python;  2）Understanding the theory of SVMs and k-Means;  3）Knowing how to use machine learning approaches to process small-size datasets. | | | | | | | | | | |
| **二、实验环境**  机器硬件配置：Windows 11 24H2  软件安装：Miniconda3, Python 3.11, MATLAB R2021b | | | | | | | | | | |
| **三、实验原理及内容**  **1. 使用MATLAB验证SVM**    **2. 使用Python验证SVM**  （1）iris数据集结果：  iris = datasets.load\_iris()  X = iris.data[:, :2]  # 特征  Y = iris.target  # 标签    图1 使用iris数据集验证SVM的结果  （2）改用wine数据集结果：  wine = datasets.load\_wine()  X = wine.data[:, :2]  # 特征  Y = wine.target  # 标签    图2 使用wine数据集验证SVM的结果  （3）iris数据集，修改RBF中的gamma数值（下图分别为0.2, 0.4, 0.7, 1）  rbf\_svc = svm.SVC(kernel='rbf', gamma=0.7, C=C).fit(X, Y)      图3 不同的gamma对RBF核SVM结果的变化  （4）iris数据集，修改Poly中的degree数值以修改多项式核函数的度数，使用更高阶数的多项式。下图多项式的阶数依次为2，3，4，5  poly\_svc = svm.SVC(kernel='poly', degree=3, C=C).fit(X, Y)      图4 不同的degree数值对Poly核SVM的结果变化  **3. 使用MATLAB验证K-means**    图5 使用MATLAB验证K-means的原始数据分布图    图6 使用MATLAB验证K-means的结果图  **4. 使用Python验证K-means**  （1）iris数据集结果：  iris = datasets.load\_iris()  X = iris.data  # 特征  y = iris.target  # 标签      图7 使用iris数据集验证K-means的结果  （2）wine数据集结果：  wine = datasets.load\_wine()  X = wine.data  # 特征  y = wine.target  # 标签      图8 使用wine数据集验证K-means的结果  （3）iris数据集，不同的n\_init数值比较  estimators = [('k\_means\_init\_1', KMeans(n\_clusters=3, n\_init=1, init='random')),                ('k\_means\_init\_10', KMeans(n\_clusters=3, n\_init=10, init='random')),                ('k\_means\_init\_20', KMeans(n\_clusters=3, n\_init=20, init='random'))]    图9 不同的n\_init数值对结果的影响  （4）设计一种算法，制定聚类评价标准  采用**轮廓系数，Calinski-Harabasz指数，Davies-Bouldin指数**来评估每个聚类结果的质量，并根据此指标选择最佳的聚类。  # Calculate evaluation metrics for the current clustering  silhouette = silhouette\_score(X, labels)  # Silhouette Score  ch\_score = calinski\_harabasz\_score(X, labels)  # Calinski-Harabasz Index  db\_score = davies\_bouldin\_score(X, labels)  # Davies-Bouldin Index  metrics.append({      'name': name,      'silhouette\_score': silhouette,      'calinski\_harabasz\_score': ch\_score,      'davies\_bouldin\_score': db\_score  })  # Print the metrics for each clustering  for metric in metrics:      print(f"Clustering: {metric['name']}")      print(f"Silhouette Score: {metric['silhouette\_score']:.4f}")      print(f"Calinski-Harabasz Score: {metric['calinski\_harabasz\_score']:.4f}")      print(f"Davies-Bouldin Score: {metric['davies\_bouldin\_score']:.4f}")      print()  最终评判结果如下：    图10 聚类评价结果图  轮廓系数（Silhouette Score）：值越大越好，大于0的值表示聚类质量较好。  Calinski-Harabasz 指数：值越大越好，表示簇之间的分离度更大。  Davies-Bouldin 指数：值越小越好，表示簇之间的相似度更小。 | | | | | | | | | | |
| **四、实验小结**  本次实验旨在帮助我熟悉使用MATLAB和Python进行机器学习算法的基本操作，重点是支持向量机（SVM）和K-means聚类算法的实现与应用。通过这次实验，我不仅掌握了如何在MATLAB和Python中实现这些常见的机器学习方法，还加深了对SVM和K-means算法理论的理解，并且学会了如何使用不同的评价标准来评估聚类结果。  在SVM部分，使用了经典的Iris数据集，通过SVM模型进行分类。我尝试了不同的核函数（线性核、RBF核、多项式核）及其超参数调整，观察了不同参数对分类结果的影响。  在K-means聚类部分，使用了Iris和Wine数据集进行聚类分析，并测试了不同的聚类数和初始化策略对结果的影响。在调整n\_init参数时，我发现随着初始化次数的增加，K-means算法的稳定性有所提升，聚类结果的质量也逐渐改善。我通过计算评价指标，对不同的聚类结果进行了量化评估。这些评价指标帮助我客观地分析和对比不同聚类模型的优劣，最终通过综合分析选择出最优的聚类结果。  通过实验，我还进一步了解了如何使用不同的机器学习数据集验证这些算法。除了经典的Iris数据集，我还使用了Wine数据集进行相同的算法验证，进一步验证了算法在不同数据集上的适应性和效果。实验还促使我思考了其他潜在的聚类算法，并鼓励我进行算法对比分析，探索不同算法在不同场景下的优势与不足。  本次实验不仅提升了我在MATLAB和Python环境中实现机器学习算法的能力，还让我理解了聚类算法的基本思想和评价标准。通过实验中的不断调试和分析，我积累了宝贵的实践经验，为今后在实际问题中应用机器学习方法打下了坚实的基础。 | | | | | | | | | | | |
| **五、指导教师评语** | | | | | | | | | | | |
| **成 绩** |  | | **批阅人** | |  | | | **日 期** |  | | |