1、根据逻辑回归分类器的预测分数对所有样本排序，并按此顺序逐个把样本作为正类进行预测，以“假正例率（FPR）”为横轴，“真正例率（TPR）” 为纵轴，可得到ROC曲线。已知FPR=FP/N，TPR=TP/P，其中P与N分别表示正类与负类的样本个数。请绘制ROC曲线。

|  |  |  |
| --- | --- | --- |
| 样本 | 类别 | 预测分数 |
| 1 | 正类 | 0.9 |
| 2 | 正类 | 0.8 |
| 3 | 负类 | 0.8 |
| 4 | 正类 | 0.7 |
| 5 | 正类 | 0.6 |
| 6 | 正类 | 0.5 |
| 7 | 负类 | 0.4 |
| 8 | 负类 | 0.3 |
| 9 | 负类 | 0.2 |
| 10 | 负类 | 0.1 |

2、假设交叉验证集中有m=1000个样本。通过逻辑回归模型预测的类别与实际类的混淆矩阵如下图所示。试求分类器的准确率（accuracy）、精度（precision）、召回率（recall）以及F1分数是多少？

|  |  |  |
| --- | --- | --- |
|  | 正类 | 负类 |
| 预测为正类 | 80(TP) | 300(FP) |
| 预测为负类 | 120(FN) | 500(TN) |

3、公式推导：**最小二乘法、多元线性回归与岭回归、逻辑回归（极大似然法）**