类别 1：（4，3）
类别 2：（4，2）
两种划分各自的信息量为：$H(1)=4\log_2 4-2\log_2 2-2\log_2 2=4\log_2 2=4$，$H(2)=2\log_2 2-\log_2 1-\log_2 1=2\log_2 2=2$
信息增益为：$IG=H(P)-H(1)-H(2)=0$
（4 分，只需要给出任意一个结果即可）第二维特征有 1、2、3、4 三个取值，因此分界点可取在 1、2 之间或 2、3 之间或 3、4 之间
（1）若分界点取在 1、2 之间，此时得到两种划分：
类别 1：（2，1）
和
类别 1：（2，2），（4，3）
类别 2：（2，3），（2，4），（4，2）
两种划分各自的信息量为：$H(1)=\log_2 1-\log_2 1=0$，$H(2)=5\log_2 5-2\log_2 2-3\log_2 3$
信息增益为：$IG=H(P)-H(1)-H(2)=8\log_2 2-5\log_2 5+3\log_2 3>0$
（2）若分界点取在 2、3 之间，此时得到两种划分：
类别 1：（2，1），（2，2）
类别 2：（4，2）
和
类别 1：（4，3）
类别 2：（2，3），（2，4）
两种划分各自的信息量为：$H(1)=3\log_2 3-2\log_2 2-\log_2 1$，$H(2)=3\log_2 3-\log_2 1-2\log_2 2$
信息增益为：$IG=H(P)-H(1)-H(2)=10\log_2 2-6\log_2 3>0$
（3）若分界点取在 3、4 之间，此时得到两种划分：
类别 1：（2，1），（2，2），（4，3）
类别 2：（2，3），（4，2）
和
类别 1：
类别 2：（2，4）
两种划分各自的信息量为：$H(1)=5\log_2 5-2\log_2 2-3\log_2 3$，$H(2)=\log_2 1-\log_2 1=0$
信息增益为：$IG=H(P)-H(1)-H(2)=8\log_2 2-5\log_2 5+3\log_2 3>0$
（5 分）因此，采用第二维特征进行分类可以获得最大的信息增益。