# 数据挖掘作业二报告

## 关联规则挖掘

**姓名：赵赫 学号：2120171103**

1. **问题描述**

本次作业中，我选择将对NFL Play-by-Play 2009-2017数据集进行关联规则挖掘。

本次关联规则挖掘任务包含以下四个子任务：

* 对数据集进行处理，转换成适合关联规则挖掘的形式。
* 找出频繁项集。
* 导出关联规则，计算其支持度和置信度。
* 对规则进行评价，可使用Lift，也可以使用教材中所提及的其他指标。

1. **数据描述**

* **数据集1：NFL Play-by-Play 2009-2017**

该数据集共包含101个属性，407656条数据记录。

1. **关联规则挖掘过程**

**3.1 将数据集转换成适合关联规则挖掘的形式**

* **数据摘要：**
* **数据可视化：**

**3.2 找出频繁项集**

* **可视化对比：**

**3.3 导出关联规则并计算支持度、置信度、Lift指标**

**3.4 挖掘结果及分析**