在机器学习领域，有一句话说“数据和特征决定了机器学习的上限，而模型和算法只是逼近这个上限而已。”可见特征工程和数据在机器学习中占有相当重要的位置。在实际的应用中，很多数据挖掘比赛例如天池比赛，优秀选手最后的成绩往往是特征做做足了文章从而使得最后的效果比较好，因此数据是基础，特征是成功的关键。











