Predicting Minimum Clicks in Minesweeper

Yan Haoyang
yanhy@shanghaitech.edu.cn

Abstract

In this article, we introduce some efforts on esimate the minimum clicks in Minesweeper. We use multiple linear regression and naive bayes classifier to predict the minimum clicks from known data.

1 Backgroud

Minesweeper is a single-player puzzle vedeo game originates from 1960s. After published on Windows 3.1 in 1992, it becomes one of the most popular games in the world. The rule of Minesweeper is simple. There are several mines in the board, and each non-mine cell has a number which indicates the amount of mines surrounding it, min on 0 (not display) and max on 8. Each leftclick can flip a non-mine cell to number (if clicks on 0, it will open a continuous part until the boundary is not 0), each rightclick can lable a mine and each doubleclick on the flipped number with all surrounding mines labled can flip all the other surrounding unflipped cells. The goal of Minesweeper is to flip all the non-mine cell without flipping any mine.

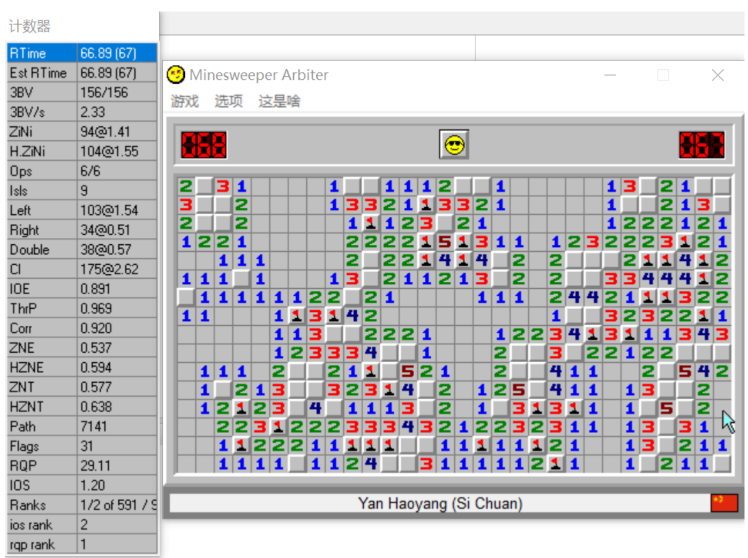

Figure 1: Minesweeper

In the expert mode of Minesweeper, the board has 30 * 18 = 480 in total with 99 mines randomly distributed in it, which means it has an enormous number of C_{480}^{99} possible boards.

In the past few decades, many competitive players of Minesweeper made a hard try to optimize the strategy and improve the speed. The recent world record made by Ju Ze-en is solving the expert mode in 28.84 seconds.

Submitted to 35th Conference on Neural Information Processing Systems (NeurIPS 2021). Do not distribute.

2 Effort on Finding the Minimum Clicks

To begin with, the basic strategy of minesweeping can be devided into 2 kinds, Flag and Non-flag. Flag means use all 3 kinds of clicks and Non-flag means only use leftclick without lable any mines. The minimum clicks of Non-flag is easy to compute while the minimum clicks of Flag is an Non-deterministic Polynomial problem which is hard to solve.

2.1 Ops, 3BVs and Isls

Ops is the short of Openings. It is an area of non-displayed 0 and all the numbers surrounding. Flipping an Op just need one leftclick.

From this, the minimum clicks of Non-flag is easy to solve. It equals to the number of Ops plus the number of all the other number-cell not in any Ops. It is named Bechtel's Board Benchmark Value, shorted in 3BVs.

Also, an island means a continuous area of non-Ops number-cell, shorted in Isls.

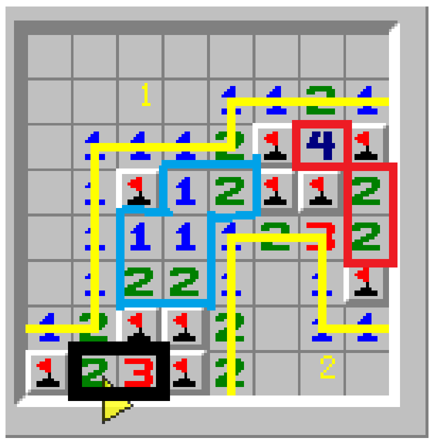

Figure 2: An example of Ops, 3BVs and Isls

An example of Ops, 3BVs and Isls is given above. This board has 2 Ops and 3 Isls with 11 non-Ops number-cell in them. Which means the minimum left-click-only clicks, that is 3BVs, of this board, is 13.

2.2 GZiNi and HZiNi

However, the problem of minimum clicks of Flag, or generally speaking, the Minesweeper, is hard and cannot be solved in polynomial time.

Intuitively, the more Ops and Isls the board have, the more seperately the position of 3BVs distribute. Which means we cannot open many 3BVs by a single doubleclick. This guide us to Ops and Isls may have a negative correlation with the minimum clicks.

In 2009, Elmar Zimmermann and Christoph Nikolaus developed two heuristic algorithm to estimate the upper bound of the minimum clicks, the Greedy ZiNi and the Human ZiNi.

The basic steps of ZiNi is to search all the area of the board and found the place where labeling mine and doubleclick has the highest return. For example, labeling one mine and then doubleclick the nearby number can flip 3 other 3BVs, the return of this step is 3-2=1. When the maximum return is larger than 0 we excute it and continue to iterate, otherwise we just compute the 3BVs remains.

The different of GZiNi and HZiNi is that, GZiNi has a random start and we repeat it many times to find the minimum solution, while in HZiNi, we first flip all the Ops and then start to iterate.

The detailed algorithm and codes can be found in $http://minesweeper.info/forum/viewtopic.php?f=15\&t=70$

3 Predict the Minimum Clicks from Known Data

ZiNi shows an greedy way to estimate the minimum clicks. However, it needs huge number of random repeat to approach the bound and usually not end on the true minimum solution.

In this section, we use multiple linear regression and naive-bayes classfier to fit the minimum clicks with known coefficients.

We use 591 rows of data from my expert mode history, first 500 rows as training set and last 91 as validation set, with 11 coloums of input data, including 3BVs, the amount of number 1 8 in the board, Ops and Isls. We use the HZiNi and GZiNi as the object data.

In the part of multiple linear regression, we get standard error of 3.88 for HZiNi and 4.27 for GZiNi in the validation set.

In the naive-bayes classfier, we preprocess the data into a small number of class. For example, for 3BVs distributed in [105,231], we seperate it into 14 classes with each range 10. The classes' amount of 11 coloums of input data are (14,4,4,4,3,5,5,3,2,8,11), and the classes' amount of 2 coloums of output data are both 6. To avoid the new class which our training set does not have, we use the Laplacian Correction, adding 1 on both denominator and numerator.

As a result, we correctedly predict the class of 80 HZiNi and 77 GZiNi in the 91 rows of validation set.

References

[1] 2009, Elmar Zimmermann & Christoph Nikolaus $http://minesweeper.info/forum/viewtopic.php?f=15\&t=70$