作業 1\_資料探勘預處理實作-第 5 組

B10801016 李亦婷

B10801018 宋 揚

B10801019 黃湘宜

一、資料集說明

資料比數: 3196

特徵數量: 37

Missing Data: 0

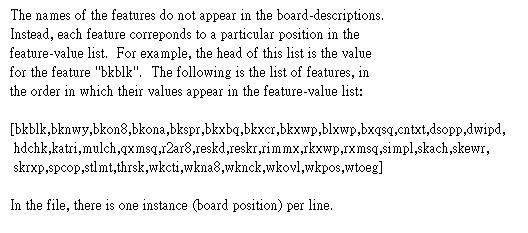
特徵資訊:

|  |  |  |
| --- | --- | --- |
| BK | BLACK KING | 黑色國王 |
| BR | BLACK ROOT | 黑色城堡 |
| WK | WHITE KING | 白色國王 |
| WP | WHITE PAWN | 白色小兵 |

|  |  |  |
| --- | --- | --- |
| **特徵名稱** | **中文特徵名** | **資料型態** |
| bkblk | BK不在路上 | boolean |
| bknwy | BK不在BR路上 | boolean |
| bkon8 | 第8行的BK來支援BR | boolean |
| bkona | A列的BK來支援BR | boolean |
| bkspr | BK可以支援BR | boolean |
| bkxbq | BK不會被升級的WP攻擊 | boolean |
| bkxcr | BK可以攻擊b7 | boolean |
| bkxwp | BK可以攻擊WP | boolean |
| blxwp | B攻擊WP | boolean |
| bxqsq | 一或以上黑色棋子控制升級的方塊 | boolean |
| cntxt | WK在一個邊緣但不在a8 | boolean |
| dsopp | 兩個國王在正常的位置 | boolean |
| dwipd | WK離交叉點很遠 | boolean |
| hdchk | 有一個隱藏勝利點所以時間延長 | boolean |
| katri | BK控制著交叉點 | boolean |
| mulch | B可以扭轉局面 | boolean |
| qxmsq | 將死的位置被升級的WP攻擊 | boolean |
| r2ar8 | BR沒辦法安全進入列A或行8 | boolean |
| reskd | WK可以被一個延遲的skewer攻擊 | boolean |
| reskr | BR自己可以造成skewer威脅 | boolean |
| rimmx | BR可以安全的被抓到 | boolean |
| rkxwp | BR對WP有影響 | boolean |
| rxmsq | BR可以安全地攻擊將死位置 | boolean |
| simpl | 一個簡單的模式 | boolean |
| skach | WK可以在一次或多次威脅到國王後被skewer | boolean |
| skewr | 在對向交叉有一個潛在的skewer | boolean |
| skrxp | BR可以skewer或BK攻擊WP | boolean |
| spcop | 有一個特殊的格局 | boolean |
| stlmt | Wk在僵局 | boolean |
| thrsk | 有一個skewer威脅潛伏著 | boolean |
| wkcti | WK不能攻擊交叉點 | boolean |
| wkna8 | WK在a8上 | boolean |
| wknck | WK被威脅 | boolean |
| wkovl | WK超負荷 | boolean |
| wkpos | WK在一個潛在的skewer點 | boolean |
| wtoeg | WK離相關邊緣只隔一格 | boolean |
| "win" or "nowin" | 比賽結果 | boolean |

二、分類處理 (包含原因)

其他特徵資料皆以縮寫的形式分成兩或三類資料，如下圖所示



所以我們將此資料中的[bkblk,bknwy,bkon8,bkona,bkspr,bkxbq,bkxcr,bkxwp,blxwp,bxqsq,cntxt,dsopp,

hdchk, mulch,qxmsq,r2ar8,reskd,reskr,rimmx,rkxwp,rxmsq,simpl,skach,skewr,

skrxp,spcop,stlmt,thrsk,wkcti,wkna8,wknck,wkovl,wkpos]，分成TRUE與FALSE兩類，如下表所示

|  |  |
| --- | --- |
| t | TRUE |
| f | FALSE |

將此資料中的wtoeg與katri，分成BLACK、WHITE與NOT GIVEN三類，如下表所示

|  |  |
| --- | --- |
| b | BLACK |
| w | WHITE |
| n | NOT GIVEN |

將此資料中的dwipd，分成兩類，如下表所示

|  |  |
| --- | --- |
| g | GREAT |
| l | LITTLE |

三、將整理好資料輸入Weka

