

5. Rule-Based Classifier 規則分類直接方式的 **RIPPER (Direct Method: RIPPER)**

RIPPER演算法，這也是一個在直接方式中常用的演算法，屬於演繹法 (general-to-specific) 的成長策略，也是基於連續覆蓋 (Sequential Covering) 且利用FOIL's Information Gain，來挑選最好的組合。

用在二元分類問題上：先選擇一個目標類別做為「正面 (Positive) 」，其他類別就是「負面 (Negative) 」，學習的規則就依照類別Positive的紀錄，Negative記錄就當做預設類別。

用在多元分類問題上：學習的規則開始於Class數量最少的開始，會有較好的精準度，其餘的為Negative。接著，最少的選擇完後，在找第二少的開始，依此類推。

RIPPER演算法 規則的成長：

1.開始於一個空的規則 $R\{\} \rightarrow y$

2.開始加入條件，只要能夠改善 FOIL's Information Gain

3.停止條件在於條件開始覆蓋Negative的資料

4.接著立即將驗證案例導入做REP (Reduced Error Pruning) 測試

5.以下面的公式來決定規則是否要進行修剪： $V = (p-n) / (p+n)$

p是規則能夠覆蓋多少個驗證資料集的Positive資料

n是規則能夠覆蓋多少個驗證資料集的Negative資料

6.如果有所改善，則刪除規則中最後一個加入的條件

比方說有一規則 $R\{A,B,C,D\} \rightarrow y$ ，則先考慮刪除D、再來是CD、再來BCD，依此類推