

# 基於機器學習分類器之程式碼抄襲比對系統

系所／電腦與通訊工程學系

指導老師／陳建伯

組員／黃韻婷、陳依婷、魏玟捷、李佩容

本研究的目的是開發出一套利用分類器進行程式碼分類的系統。本系統是由教師端、學生端、資料庫、計算特徵程式及分類器所組成，如圖 1 所示。此系統先定義程式碼特徵，並擷取程式碼中的特徵值，將特徵值經過正規化處理，再透過分類器將同學所繳交的作業進行分類，分在同一類的作業，即是有可能互相抄襲。此分類的結果，可以提供給教師作為評分的參考。

進行特徵擷取並存於資料庫。再透過分類器進行分類，分類的結果即為程式碼抄襲的判斷依據。

最後將會顯示各學生特徵數量、相似度與程式是否編譯成功及分類結果的分類頁面，如圖 2 所示，若教師想檢查學生程式碼，只需在分類結果點選學生學號即可，如圖 3 所示。

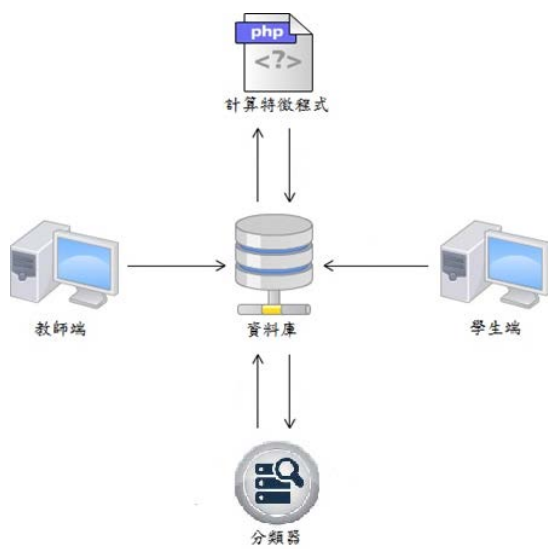


圖 1：系統流程圖

當教師創建作業完畢後，待大部分學生都繳交作業後，教師即可進行作業分析。系統將會依據教師所選擇的作業

The screenshot shows a web browser displaying a table titled '學生相似度比較表' (Student Similarity Comparison Table). The table lists student IDs in the first column and similarity scores in the subsequent columns. The scores range from 67 to 100.

學號	00450110	00450331	00450374	00450603	00450676	00451004	01050036	01050072	01050081	01050143	01050106	01050105
00450110	94	94	94	78	78	76	68	93	83	72	78	78
00450331	94	100	100	72	72	74	62	92	92	67	73	73
00450374	94	100	100	72	72	74	62	92	92	67	73	73
00450603	94	100	100	72	72	74	62	92	92	67	73	73
00450676	78	72	72	72	100	86	67	77	77	67	79	79
00451004	78	72	72	72	100	86	67	77	77	67	79	79
01050036	78	74	74	74	80	80	75	81	81	74	86	86
01050072	68	62	62	62	67	75		64	84	63	74	74
01050081	93	92	92	92	77	77	61	64		100	69	69
01050143	93	92	92	92	77	77	61	64	100		69	69
01050106	72	67	67	67	67	67	74	83	89	89		72
01050105	78	73	73	73	79	79	66	74	80	80	72	
01050132	67	63	63	63	68	64	71	89	89	71	85	
01050105	80	83	83	83	74	74	86	74	82	82	72	86

圖 2：相似度比較圖

The screenshot shows a web browser displaying a table titled '分類結果' (Classification Results). The table lists student IDs in the first column and their corresponding classification results in the subsequent columns.

學號	分類結果
00450110	00450331 00450374 00450603 01050081 01050143 01050106 01050105
00450331	01050072 01050081 01050106 01050105
00450374	00450603 01050036 01050072
00450603	01050072 01050081 01050106 01050105
00450676	01050036 01050072 01050081 01050106 01050105
00451004	01050036 01050072 01050081 01050106 01050105
01050036	01050072 01050081 01050106 01050105
01050072	01050081 01050106 01050105
01050081	01050106 01050105
01050143	01050106 01050105
01050106	01050105
01050105	

圖 3：分類結果