1. **Cancer-miRNA Search System**
2. Final Project for the Biomedical Information Retrieval
3. 組員：方郁文、吳政達

## 問題定義

找到 Cancer 和 mi-RNA 的關係，並顯示其證據

## 目標 Cancer 及 mi-RNA

Cancer type:

1. Osteosarcoma 骨肉瘤
2. Ewing sarcoma 尤文式骨肉瘤
3. Chondrosarcoma 軟骨骨肉瘤
4. Multiple myeloma 多發性骨髓瘤
5. Rhabdomyosarcom 橫紋肌肉瘤
6. Synovial sarcoma 關節滑囊肉瘤
7. Neuroblastoma 神經母細胞瘤
8. Cholangiocarcinoma 膽管惡性腫瘤
9. Retinoblastoma 視網膜神經母細胞瘤
10. Squamous cell carcinoma 扁平細胞癌

miRNA:

1. miR-34a
2. miR-143
3. miR-22
4. miR-106a
5. miR-181a
6. miR-141
7. miR-15a
8. miR-25
9. miR-206
10. miR-1
11. miR-183
12. miR-17
13. miR-124
14. LIN28b
15. miR-26a
16. miR-21
17. let-17
18. miR-17-92
19. miR-184
20. miR-31

## 系統環境

* 作業系統 ： Ubuntu 20.04 LTS
* 程式語言 ： Python 3.7.9
* 網站架構 ： Django 3.1.1
* 資料庫： SQLite
* 讀取csv檔 ： Pandas 1.1.3
* 文件來源 : Pubmed

## 實作方法

* 由骨科醫師決定作為目標之骨骼相關癌症及mi-RNA
* 於 PubMed 將 Cancer 及 mi-RNA 配對進行搜尋，並爬下相關文章，整理完後共有 374篇文件
* 取得每篇文章之來源期刊，共有 175 個期刊
* 人工比對找出175個期刊對應之impact factor
* 將爬取下來的文件存入資料庫
* 資料庫有五個 table
  + 文章
    - 標題 ( title )
    - 大綱 ( abstract )
    - 來源期刊 ( journal )
    - Impact factor
  + Mi-RNA
  + Cancer
  + 句子
    - 句子內容
    - 對應文件
    - 包含 miRNA
    - 包含 Cancer
    - 包含 Expression
  + Expression
    - promoted/promotion/promoter
    - higher expression/expressing/overexpression/overexpressed
    - lower/poor expression
    - regulates/regulation/up-regulated/upregulated/upregulation/up-regulation
    - enhances/enhancing
    - down-regulated/downregulation/down-regulation/ underexpression
    - suppresses/suppression/suppressed/suppressor
    - repression/ repressing
    - increased
    - decreased
    - carcinogenesis
    - inhibited/inhibition/inhibitory
    - interacted/interaction
    - axis
    - mediator/mediated
    - metastasis
    - target/target gene
    - oncomir
    - oncogenes
    - markers/biomarkers
    - p53/ tumor suppressor
* 找出句子中的 Cancer 、 mi-RNA 以及 Expression，並畫底線
* 將這些句子根據來源的期刊 impact factor 進行排序
* 定義 Evidenec Score 為 key sentences x num of articles x ( avg(impact factor + 1 ) )
* 計算 Cancer 和 mi-RNA 之間 relation 的 Evidence Score
* 實作使用者介面，利用 Django 架設網站

**介面展示**



圖1. 搜尋 Cancer 及 mi-RNA

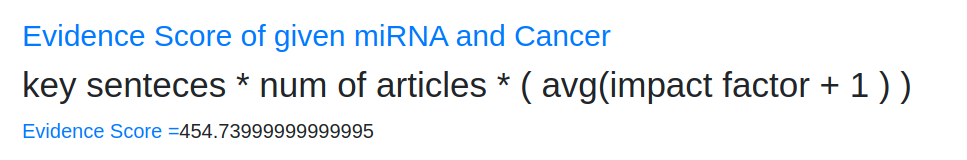
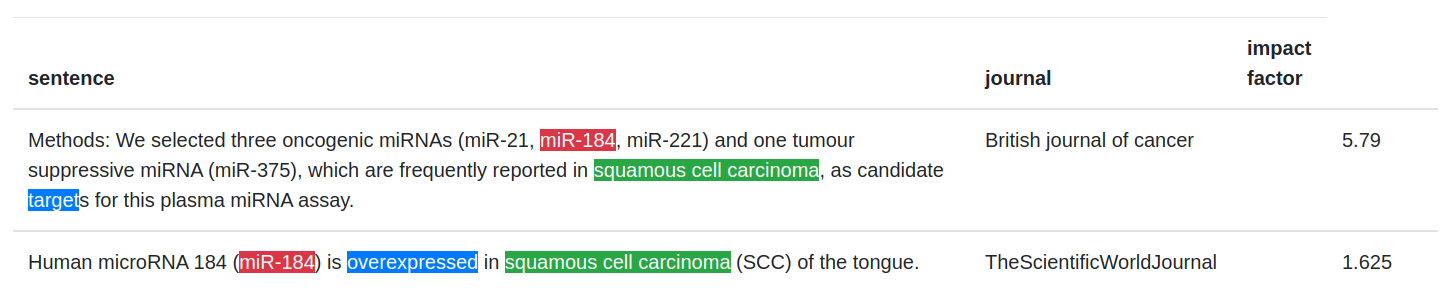
圖2. 顯示所搜尋之Cancer及mi-RNA 之關係證據分數

圖3. 標示所搜尋之Cancer及mi-RNA 之關係的表現 (Expression)



圖4. 完整介面

**分工**

政達：

* 找尋合適Cancer以及mi-RNA
* 設計Evidence Score公式
* 蒐集期刊之impact factor
* 為系統提出建議

郁文：

* 爬取 PubMed文章
* 建立並將資料輸入資料庫
* 前端介面呈現
* 實作找尋Expression及Evidence Score