Build a database of retrosynthetic chemical reactions for reinforced learning and create manuals for artificial intelligence programs such as Monte Carlo searches

By Wei-Hong Hong

Planning Division, Institute of Nuclear Energy Research

Abstract

In this study, a chemical retrosynthesis planning algorithm was programmed into an AI model that can perform chemical retrosynthesis planning for pharmaceuticals. Three types of neural networks were used to predict the possible chemical synthesis reactions and their rationality by training a large database of data from five chemical reactions. The neural networks were then used to select the plausibility of each chemical reaction step, and then a deep Monte Carlo tree search and feedback learning was performed to see if the final starting material could be found, and finally to obtain a plausible drug retrosynthesis pathway plan. A full code and operating instructions are included in this study, using the nuclear medicine TRODAT-1 as an example, to simulate and plan an innovative retrosynthesis pathway and propose a new chemical synthesis.

Keywords: chemical retrosynthesis, AI, Monte Carlo tree search, reinforcement learning

建立應用於逆合成強化學習之化學反應資料庫及蒙地卡羅搜尋等人工智慧程式之操作手冊

洪瑋嶸

核能研究所綜合計畫組

摘要

本研究將化學逆合成規畫演算法實體化，成為可執行藥物化學逆合成規劃之AI模型，利用3種神經網路經由5種化學反應數據資料庫中大筆資料的訓練，分別習得預測可能的化學合成反應及其合理性，並經由神經網路對每步化學合成反應的合理性挑選，再經過蒙地卡羅樹深度搜尋及回饋學習，看能否找得到最終的起始物基本分子，最後取得合理藥物逆合成路徑規畫。本研究附上完整的程式碼及操作說明，並以核醫藥物TRODAT-1為例，進行模擬，規劃其創新的逆合成路徑並提出新製程，最後獲得化學專家肯認，可再進一步進行實驗驗證。目前已完成化學逆合成 AI model簡易版軟體，後續可結合化學家專業知識及經驗，加速各種核醫藥物原料藥開發步驟更少、產率更高的新製程，並實際做化學驗證，證實模擬結果的可行性。

關鍵字：化學逆合成、人工智慧、蒙地卡羅搜索、強化學習。

目錄

[目錄 iii](#_Toc71532717)

[圖目錄 iv](#_Toc71532718)

[**1.用電腦工具之強化學習輔助化學逆合成之簡要說明** 1](#_Toc71532719)

[**1.1強化學習應用於蒙地卡羅之化學逆合成搜尋路徑方法** 1](#_Toc71532720)

[**1.2應用強化學習輔助化學逆合成流程說明** 2](#_Toc71532721)

[**2.資料庫建置** 4](#_Toc71532722)

[**2.1化學逆合成資料庫來源說明** 4](#_Toc71532723)

[**2.2資料庫建立條件說明** 4](#_Toc71532724)

[**2.3資料庫效度驗證說明** 5](#_Toc71532725)

[**2.4資料庫下載說明** 5](#_Toc71532726)

[**3.資料預處理** 9](#_Toc71532727)

[**3.1化學反應式轉換成各種反應規則** 9](#_Toc71532728)

[**3.2負面反應資料庫的建置** 19](#_Toc71532729)

[**3.3負面反應資料庫之修正** 21](#_Toc71532730)

[**4.神經網路訓練** 23](#_Toc71532731)

[**4.1擴展神經網路訓練** 23](#_Toc71532732)

[**4.2推演神經網路訓練** 32](#_Toc71532733)

[**4.3 In-scope filter神經網路訓練** 41](#_Toc71532734)

[**5.蒙地卡羅搜索及逆合成路徑預測** 50](#_Toc71532735)

[**6.模擬結果(以TRODAT-1為例)與總結** 67](#_Toc71532736)

[**6.1模擬結果** 67](#_Toc71532737)

[**6.2總結** 70](#_Toc71532738)

圖目錄

[圖 1強化學習應用於蒙地卡羅之化學逆合成搜尋路徑方法說明 1](#_Toc68021976)

[圖 2應用強化學習輔助化學逆合成流程說明 3](#_Toc68021977)

[圖 3化學反應資料集與eMolecules的基本分子資料集的下載方式 6](#_Toc68021978)

[圖 4基本分子資料庫建置之程式碼 8](#_Toc68021979)

[圖 5化學反應式轉換成各種反應規則之程式碼 19](#_Toc68021980)

[圖 6負面反應資料庫的建置之程式碼 21](#_Toc68021981)

[圖 7 負面反應資料庫修正之程式碼 22](#_Toc68021982)

[圖 8 擴展神經網路訓練之流程 24](#_Toc68021983)

[圖 9擴展神經網路訓練之程式碼 32](#_Toc68021984)

[圖 10推演神經網路訓練之流程 33](#_Toc68021985)

[圖 11推演神經網路訓練之程式碼 41](#_Toc68021986)

[圖 12 In-scope filter神經網路訓練之流程 42](#_Toc68021987)

[圖 13 In-scope filter神經網路訓練之程式碼 49](#_Toc68021988)

[圖 14蒙地卡羅搜索結合神經網路之程式碼 66](#_Toc68021989)

[圖 15逆合成途徑預測-以核醫藥物TRODAT-1 為例 67](#_Toc68021990)

[圖 16 TRODAT-1新製程規劃 68](#_Toc68021991)

[圖 17逆合成途徑預測-無汞製程建議測試步驟 69](#_Toc68021992)

**1.用電腦工具之強化學習輔助化學逆合成之簡要說明**

**1.1強化學習應用於蒙地卡羅之化學逆合成搜尋路徑方法**

本研究由目標化學分子為始，首先先”選擇”一條報酬分數最高之逆合成路徑，直到所有中間產物皆未找到合適的前驅物為止，此時利用”擴展神經網路”，替其一中間產物找尋及判斷合適的逆反應規則(由 in-scope filter神經網路來判斷規則是否合適)，使其可以找到合適的前驅物，再利用”推演神經網路”為其前驅物簡單快速的搜尋可能合適的逆反應規則，以期找到基本小分子可化學合成為前驅物，並回傳期望值當作此次”擴展”下的報酬分數，最後將此報酬分數”更新”至逆合成路徑的所有節點之上，之後再重新”選擇”，形成一個可無限重複的循環，但在報酬分數屢次的更新下，會強化找到成功的化學逆合成路徑的可能性，若在逆合成路徑中，所有中間產物皆為基本小分子，則可視為找到成功的化學逆合成路徑，也就是此蒙地卡羅循環找到了”終點”(如下圖)



圖 1強化學習應用於蒙地卡羅之化學逆合成搜尋路徑方法說明

**1.2應用強化學習輔助化學逆合成流程說明**

除了蒙地卡羅搜索外，還需要很多前置作業，完整的以電腦工具輔助化學逆合成流程如下圖。首先必須取得足夠多的化學反應數據，才能訓練好數個堪用的神經網路進行化學反應的預測，並結合強化學習推估化學逆合成路徑，本研究在資料庫的建置，由歷年來的專利，總共取得百萬筆數量級的資料(reaction dataset)，此外也取得兩千多萬筆可購買的化學藥品資料，作為化學合成最起始的基本小分子資料庫(basic molecules)。建置完資料庫後，必須將化學反應式中參與反應的關鍵反應中心給擷取出來，形成反應規則，以作為人工智慧學習的答案，才有辦法訓練出能推估各種化學反應的神經網路，故在資料預處理中，需將化學反應式轉換成各種反應規則(Transformation Rule Extraction)，本研究根據狹義的反應中心定義(精準、常見但不完整)及廣義的反應中心定義(完整但不精準，容易誤判)，分別找出一套規則集(Set#A及Set#B)，Set#A資料集作為” 推演神經網路(rollout)”的訓練集，Set#B資料集作為” 擴展神經網路(expansion)”的訓練集，並且運用Set#B規則資料集，由化學專家挑選不合理的規則產生不合理的產物，以形成負面反應資料庫(negative reactions)，作為判定化學反應合理性的神經網路(in-scope filter)訓練集，最後再結合上一節所述的蒙地卡羅搜索，反覆選擇、擴展、推演及更新等步驟，直到預測出特定目標分子(target molecules)的完整化學逆合成路徑(reaction path)，並經由化學實驗驗證其預測的可行性，本研究後續章節將會針對資料庫建置、資料預處理、神經網路訓練、蒙地卡羅搜索以及逆合成路徑預測還有模擬結果詳細說明，但化學實驗驗證已超出本文研究的範圍，故不詳加論述。

|  |
| --- |
|  |

圖 2應用強化學習輔助化學逆合成流程說明

**2.資料庫建置**

**2.1化學逆合成資料庫來源說明**

通過文本挖掘（text mining），由美國專利局（United States Patent and Trademark Office，USPTO）從 1976 年至 2016 年 9 月間公開的所有專利中提取化學反應。此文本挖掘方法由 Nadine Schneider（2016）發表於 J. Med. Chem.[[1]](#footnote-1)，其程式碼及化學反應資料庫已開源，可免費下載及使用

**2.2資料庫建立條件說明**

在資料處理方面，本研究有剃除重複的資料，並將化學反應資料庫經由擴展性連通指紋法（extended-connectivity fingerprints，ECFP）之演算法，轉換資料格式，建立分子結構指紋資料庫和化學反應結構指紋資料庫，而後將化學反應資料庫經由不合理規則配對演算法，從Set#B資料集中挑出不合理規則來產生不合理產物，建立負面反應資料庫，詳細的程式碼，請見資料預處理章節。

ZINC數據庫是專門為虛擬篩選準備的市售化合物(基本小分子)的精選集合。ZINC由位於加利福尼亞大學舊金山分校藥物化學系的Irwin和Shoichet實驗室提供。本研究將ZINC 資料庫中有庫存的化學藥品過濾出來，建立基本分子資料庫，此外本研究也將eMolecules(化學品販售公司)中所販賣的building products加入基本分子資料庫，以充實資料庫內容。

本研究用到的5種數據資料庫之功用說明如下:

1. 化學反應資料庫:提供AI模型大量的反應知識，使其能有效提供化學反應諮詢與建議。

2. 負面化學反應資料庫:可防止AI預測出不合理的建議反應。

3. 分子結構指紋資料庫:使AI能迅速且準確地掌握分子(例如TRODAT-1)的化學逆合成拆解方向。

4. 化學反應結構指紋資料庫:使AI能迅速且準確地給予規劃製程中每步驟適合的反應建議。

5. 起始物基本分子資料庫:使AI能拆解到最簡單且可容易購買的分子作為起始物，例如對TRODAT-1作化學逆合成拆解，可得到非鹽酸古柯鹼分子作為起始物，避免購買管制藥品的繁複流程。

**2.3資料庫效度驗證說明**

化學反應資料庫中的資料，本研究進行抽樣並由商業軟體Scifinder驗證其正確性，而負反應資料庫中的資料，抽樣由化學專家評估其不合理性。

**2.4資料庫下載說明**

下圖是化學反應資料集以及eMolecules的基本分子資料集的下載程式碼，若連結失效的話可至官方網站下載，網址分別如下: <https://figshare.com/articles/dataset/Chemical_reactions_from_US_patents_1976-Sep2016_/5104873>以及<https://downloads.emolecules.com/orderbb/>，至於Zinc的基本分子資料集可至<https://zinc.docking.org/tranches/home/>，購買purch狀態選擇agent而下載得到.txt檔，最後在命令列對該檔案用wget指令即可得到完整的Zinc基本分子資料集。

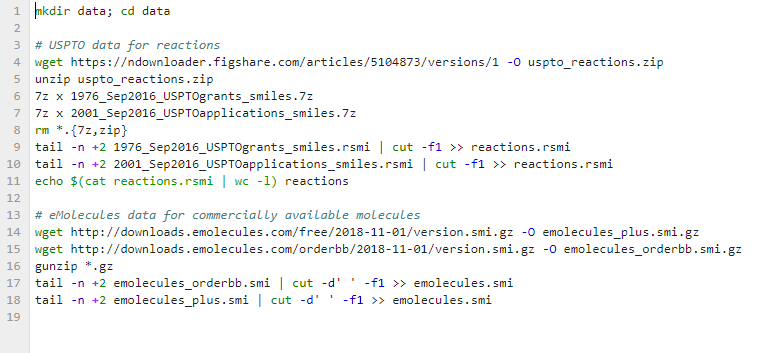


圖 3化學反應資料集與eMolecules的基本分子資料集的下載方式

下圖是eMolecules的基本分子資料庫與Zinc基本分子資料庫建置的程式碼，主要是從資料集中將本研究需要的簡化分子線性輸入規範(Simplified molecular input line entry specification, SMILES)化學分子表示式擷取出來並彙整成資料庫。





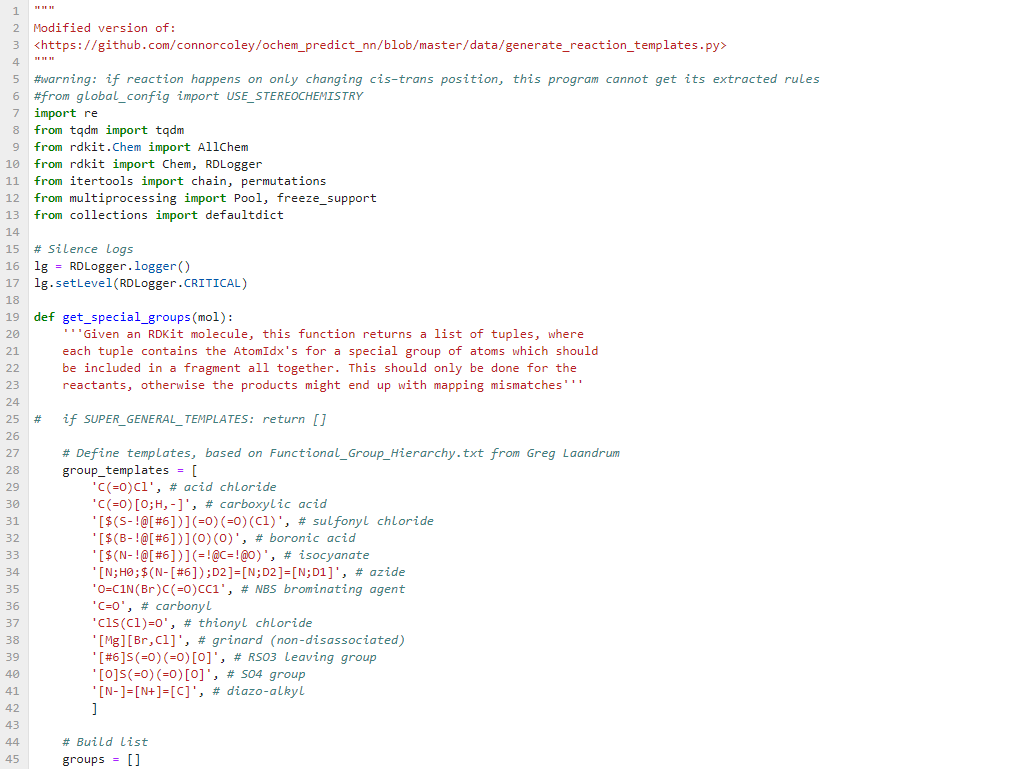


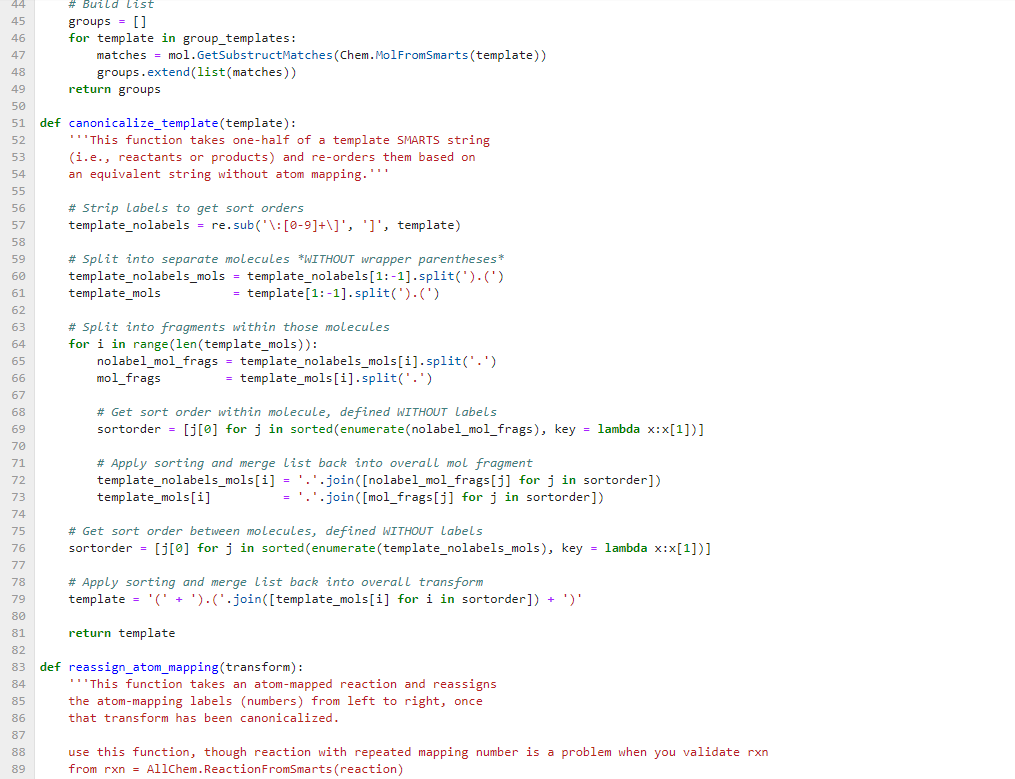
圖 4基本分子資料庫建置之程式碼

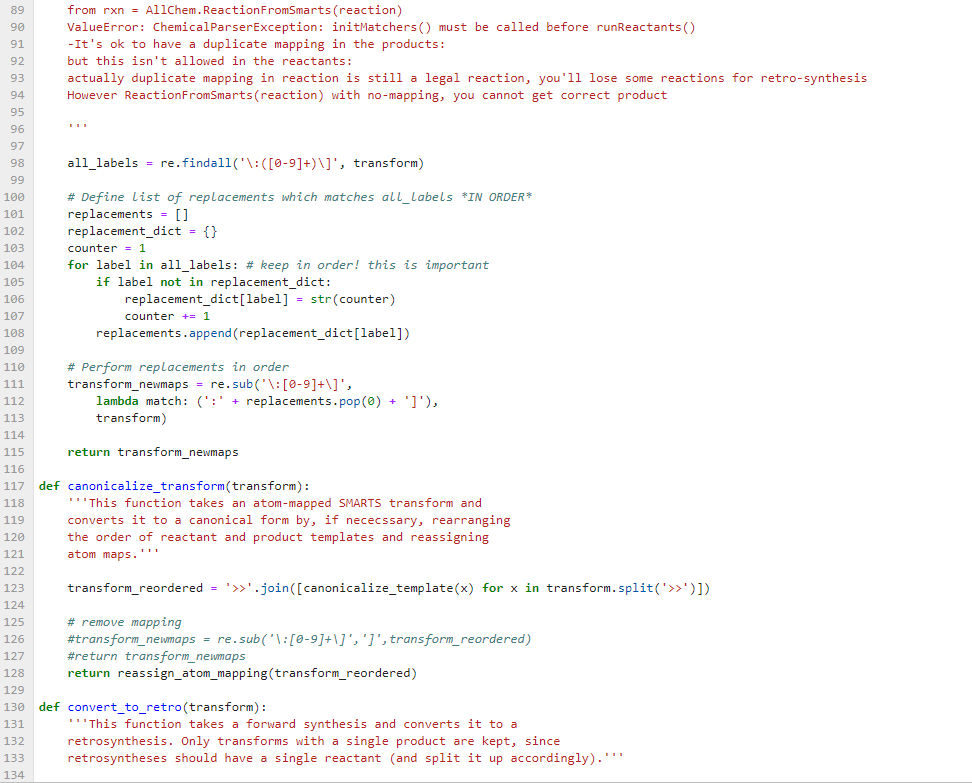
**3.資料預處理**

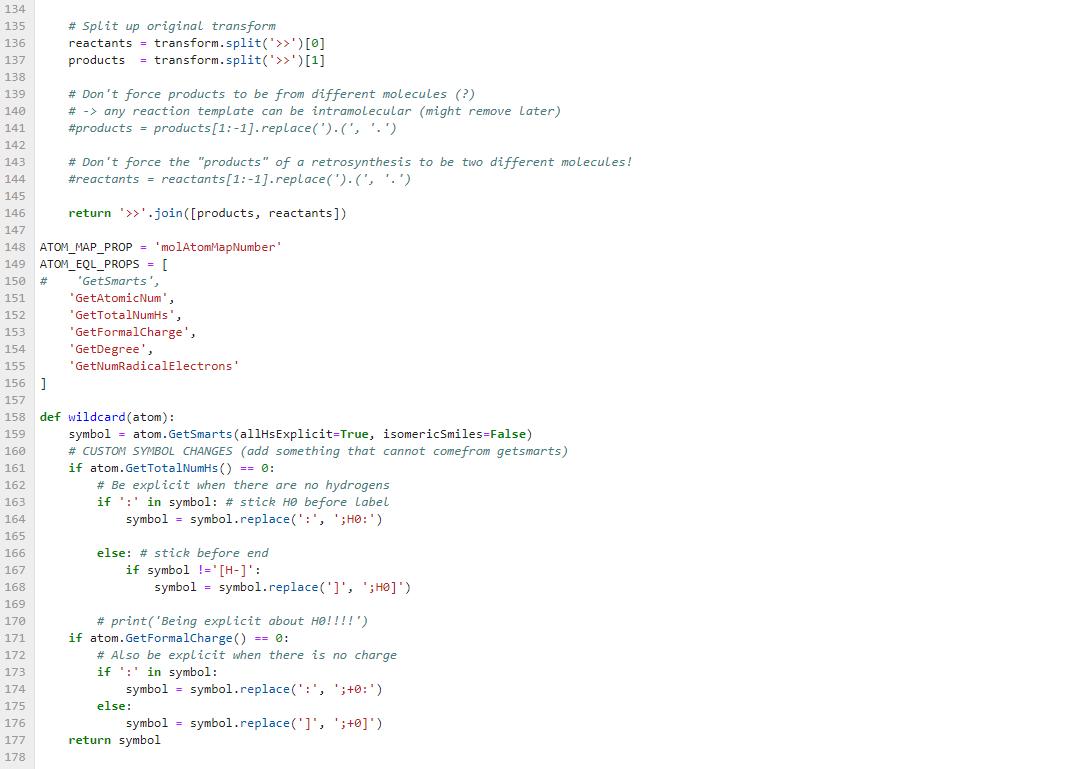
**3.1化學反應式轉換成各種反應規則**

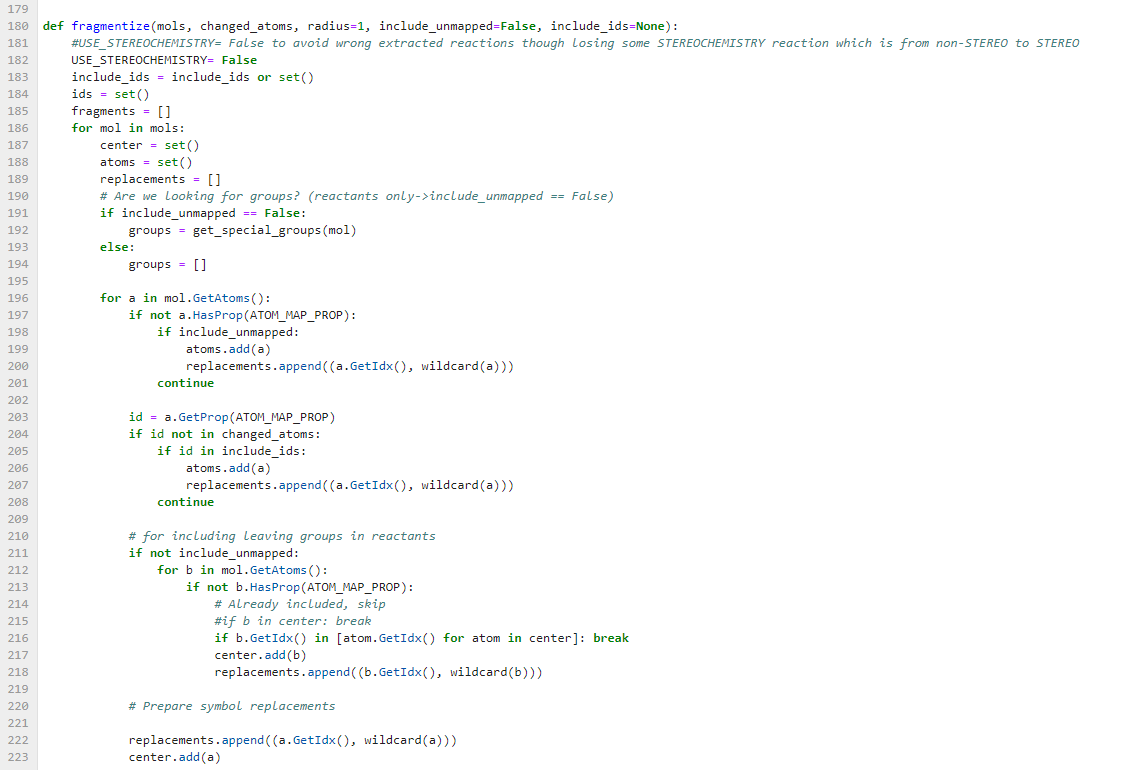
完整程式碼請見下圖，簡要說明如下，透過extract()這個函式，將化學反應資料庫裡面的資料逐一找出分別對應的反應規則，每個反應規則皆有其反應中心，本研究中狹義的反應中心定義為反應前後，同一個原子連結氫原子的數量、帶的電荷量或與周圍原子的鍵結數有發生改變時，該原子及其周圍一個原子範圍的所有原子(包含其連結氫原子數量及帶電量)都視為反應中心，還有反應物的離去基也視為反應中心；而廣義的反應中心定義為反應前後，同一個原子連結氫原子的數量、帶的電荷量或與周圍原子的鍵結數有發生改變時，該原子(包含其連結氫原子數量及帶電量)以及離去基就視為反應中心，因此廣義的反應中心規則，同樣的規則，可包含較多元的反應類型，比較不會因為資料庫沒有數據而遺漏，但是較不精準，不是所有符合同樣規則的反應物都真的會發生化學反應，之所以需要兩套不同定義的反應規則，是為了因應後續”擴展神經網路”以及”推演神經網路”不同需求所生，在”推演神經網路”中，為了快速取得報酬分數，沒辦法花時間在檢驗反應規則是否正確，所以需要比較精準且常見的反應規則集Set#A但犧牲化學逆合成的多元可能性；而在”擴展神經網路”中，是希望能預測出更多可能的不同化學逆合成途徑，所以需要一套能涵蓋廣泛化學反應的規則集Set#B，至於反應正確性與否就交由” in-scope filter 神經網路”來進行評估。在本研究中，同樣的反應規則在資料庫中出現4次以上，該反應規則才放入Set#B，而Set#A是指收錄出現15次以上的反應規則。此外為了確保擷取的反應規則是正確的，將透過val\_rc2()這個函式讓反應物與反應規則結合，確認得到的產物是否正確，也會將產物與反應規則結合，確認得到反應物的正確性，經過雙重驗證，以確保反應規則的正確。其他輔助函式的詳細說明，請參見程式碼中的「#」或「’’’」，這兩者後面接的文字內容皆為解釋各段程式碼或函式的功用。

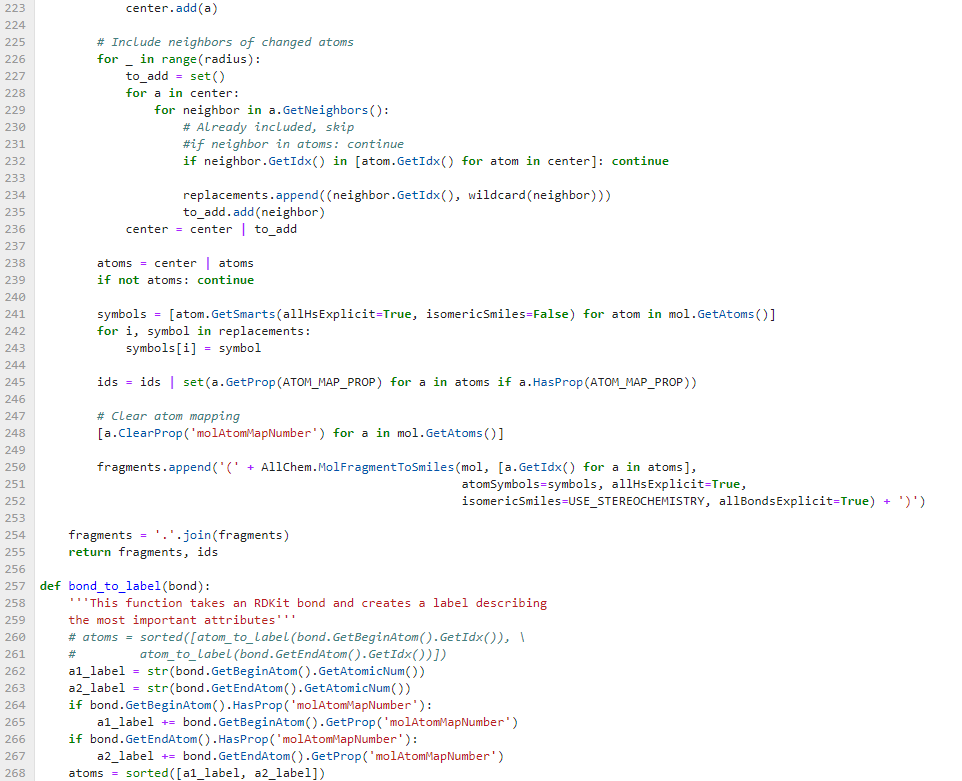


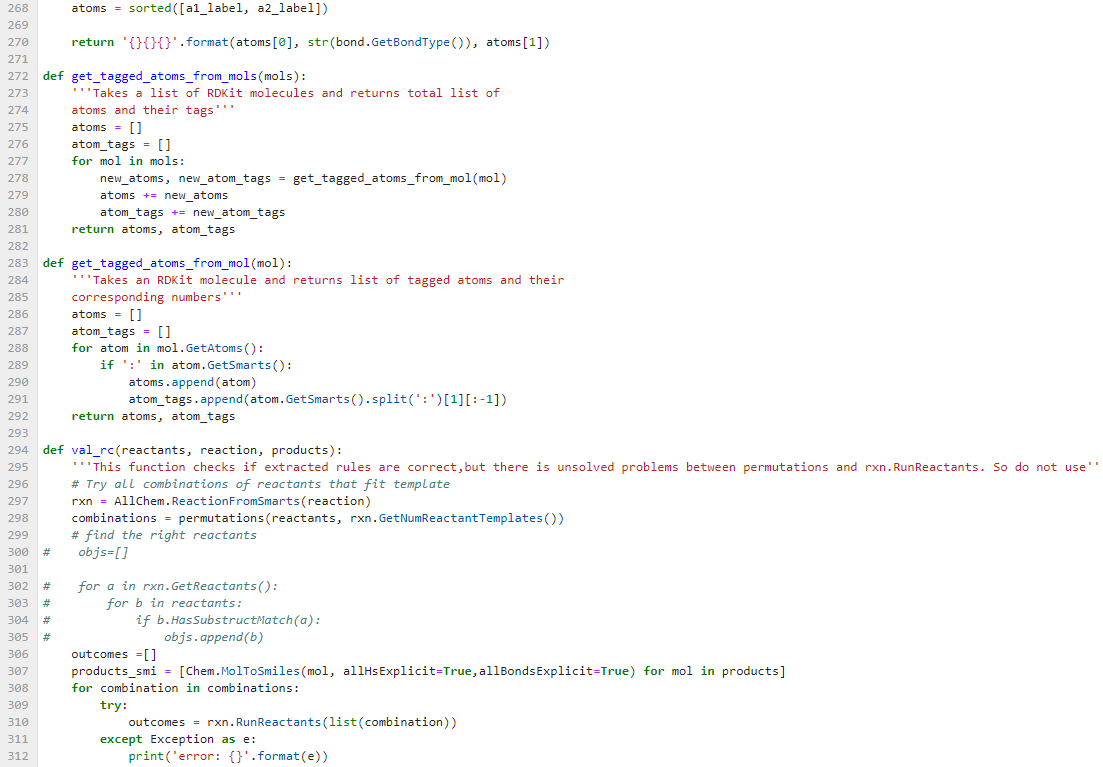








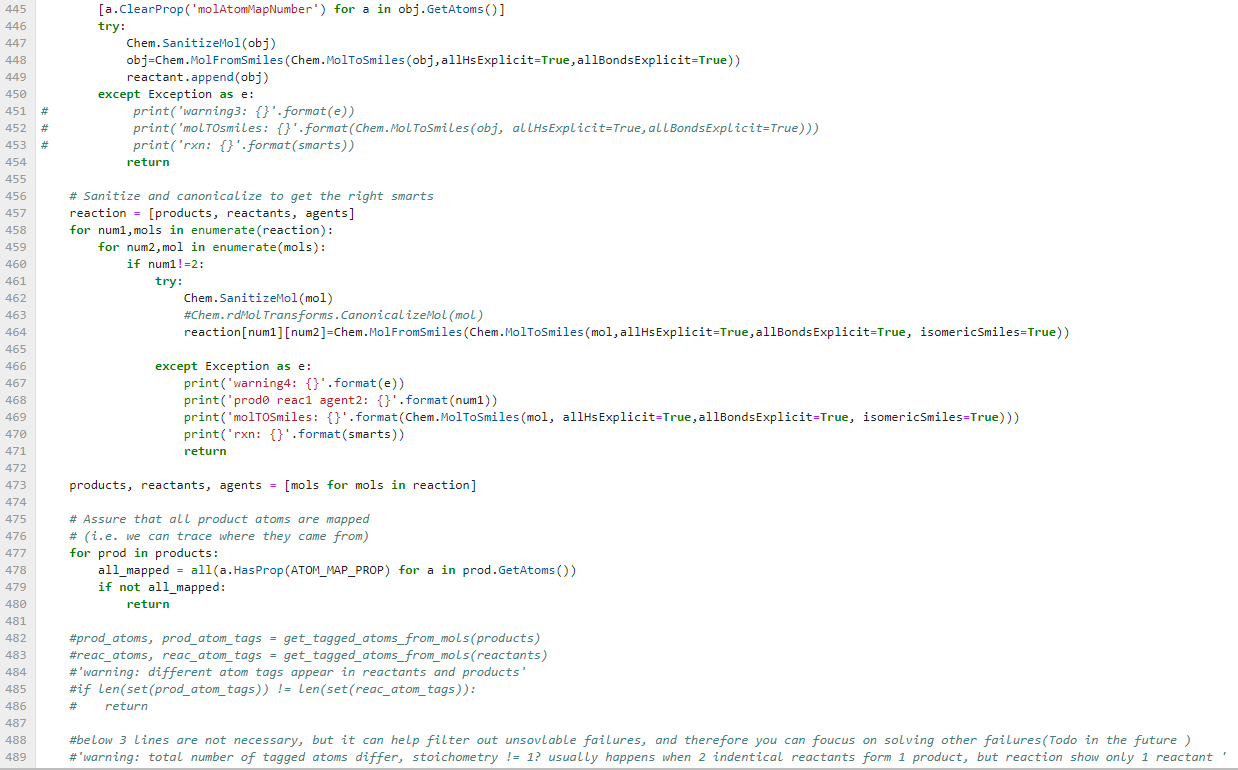


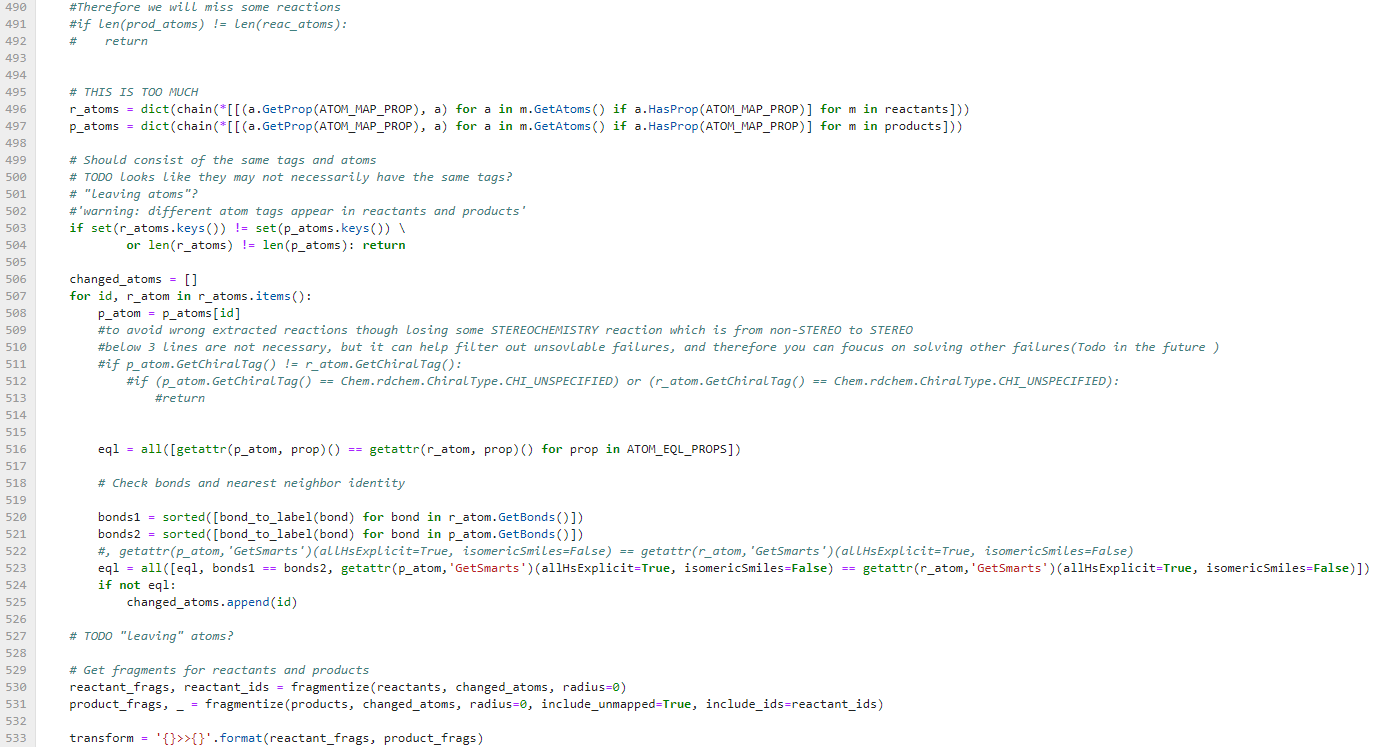
















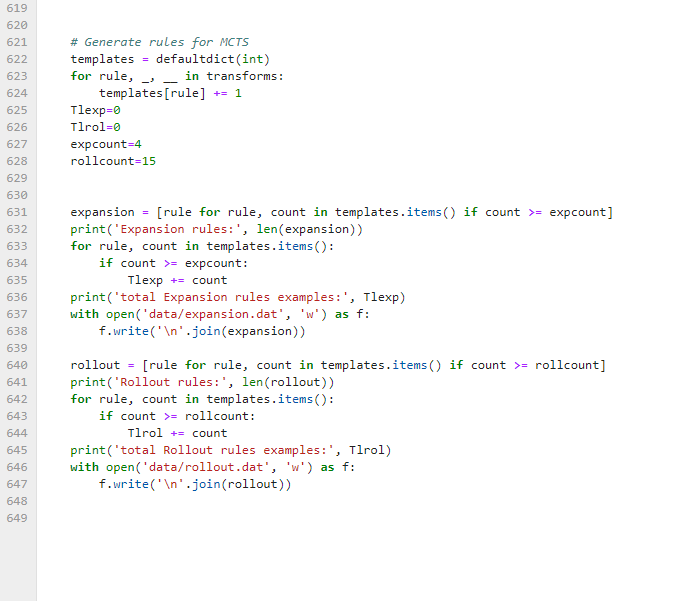
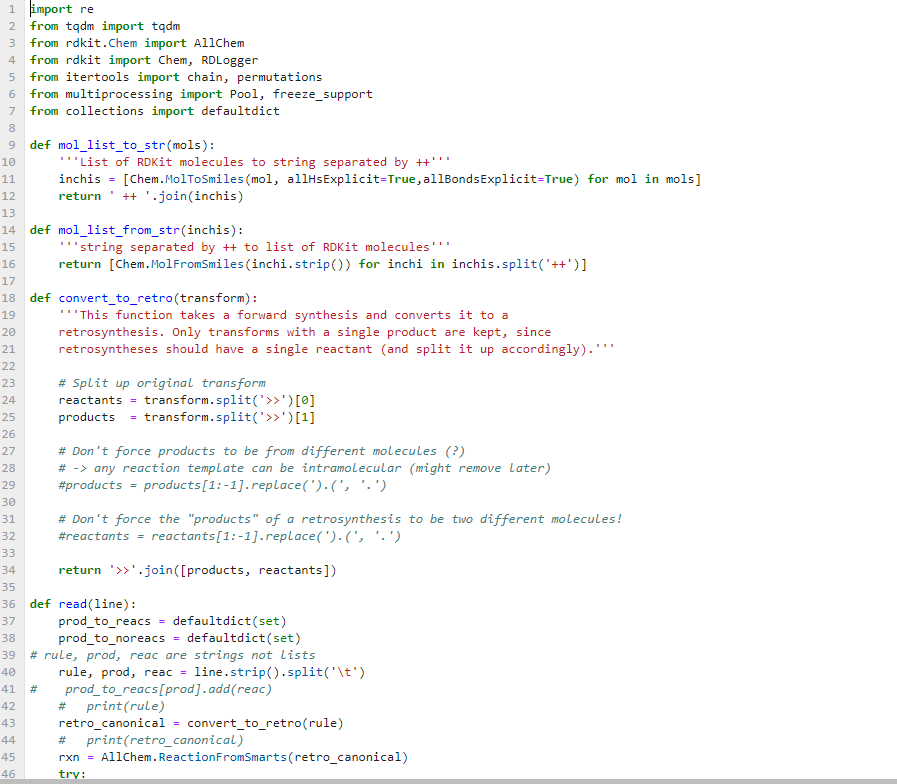
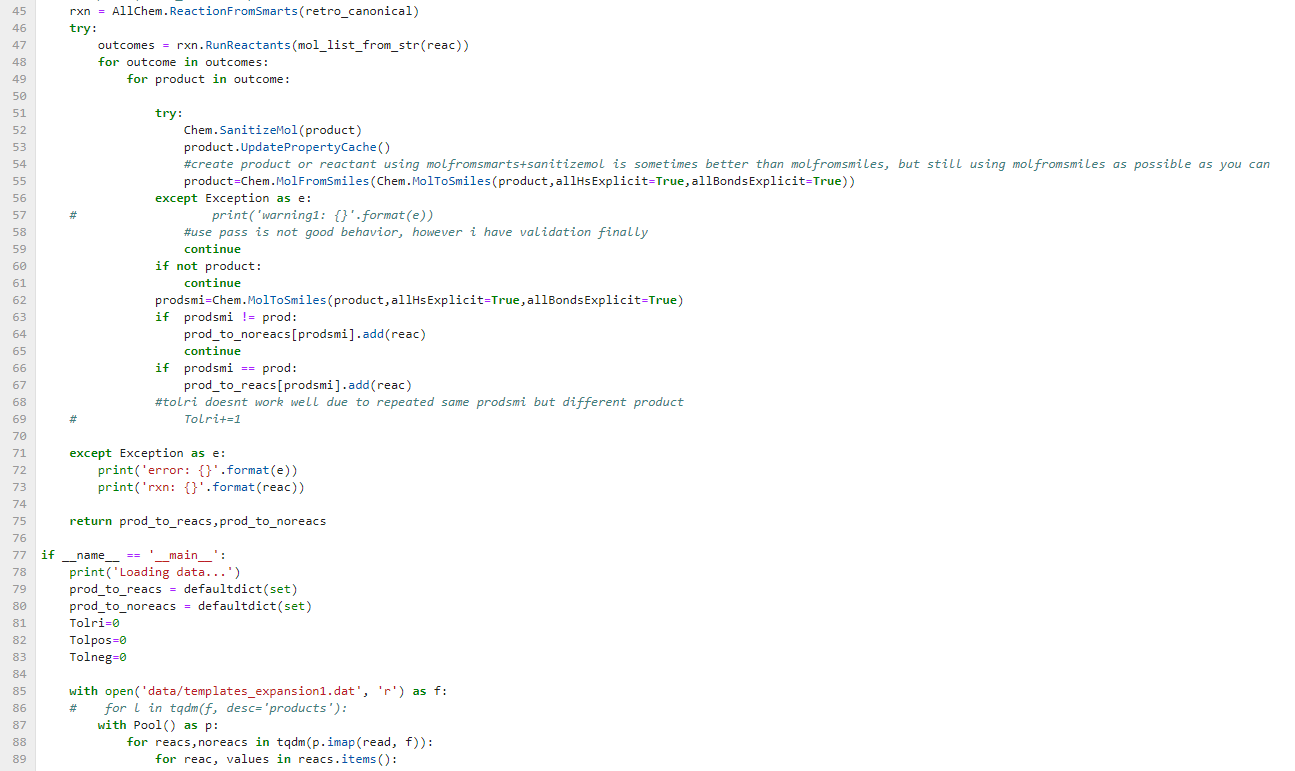


圖 5化學反應式轉換成各種反應規則之程式碼

**3.2負面反應資料庫的建置**

完整程式碼請見下圖，簡要說明如下，利用上一節擷取出來的反應的規則集Set#B，再配合函式convert\_to\_retro()將逆反應規則集轉換成正反應規則集，並將反應物與相對應的正反應規則集做搭配，若出現文獻中正確的產物，則忽略不計，但若出現其他非文獻實驗的產物，也就是正確產物的同分異構物，經化學專家判定，此類同分異構物皆為錯誤的產物，將上述反應物及錯誤的產物集中在一起，即形成一筆負面反應，反覆測試資料庫中的所有反應物，即可累積足夠量的負面反應，形成負面反應資料庫





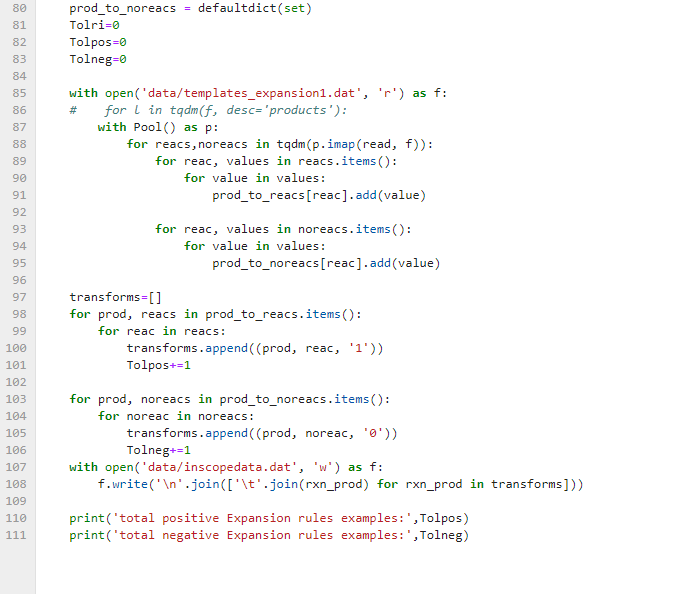


圖 6負面反應資料庫的建置之程式碼

**3.3負面反應資料庫之修正**

完整程式碼請見下圖，簡要說明如下，因為反應的規則集Set#B，是比較廣泛而不精確的反應規則，會出現同樣的反應物在不同反應條件下，產生不同的產物，但仍歸屬於同樣的反應規則，故上一節的程式碼，會將合理的同分異構物產物，誤判為錯誤的產物，故須將錯誤的產物與所有正確的產物做交互比對，以過濾掉錯誤的負面反應，完成修正。



圖 7 負面反應資料庫修正之程式碼

**4.神經網路訓練**

**4.1擴展神經網路訓練**

簡要說明如下，先用函式getmorganfingerprintasbitvect()將化學反應資料庫中各筆資料的產物分子指紋擷取出來，當作神經網路的輸入資料，反應的規則集Set#B當成答案的選項，並把產物分子對應的反應規則當成正確答案，來做神經網路權重的訓練。神經網路的第一層為512個神經元的dense層，並經過ELU激活函數，而後經由Dropout層，以防神經網路過擬合，而後反覆連結Highway層(具有可調的線性及非線性函數神經元之優點)及Dropout層五次，最後透過Softmax層，找出各種可能反應規則的機率分布，下圖為流程示意圖



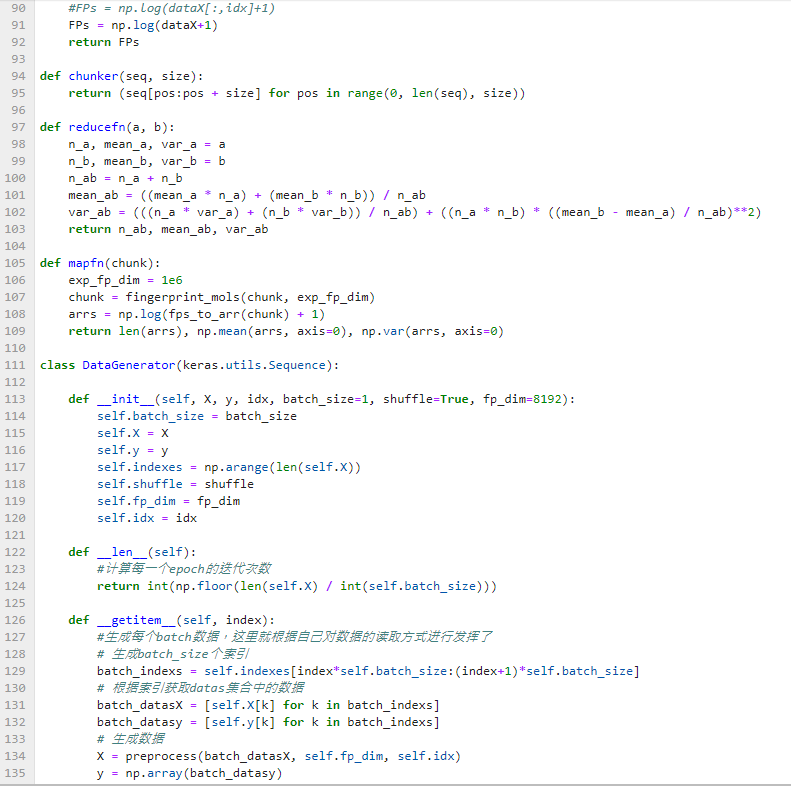
圖 8 擴展神經網路訓練之流程

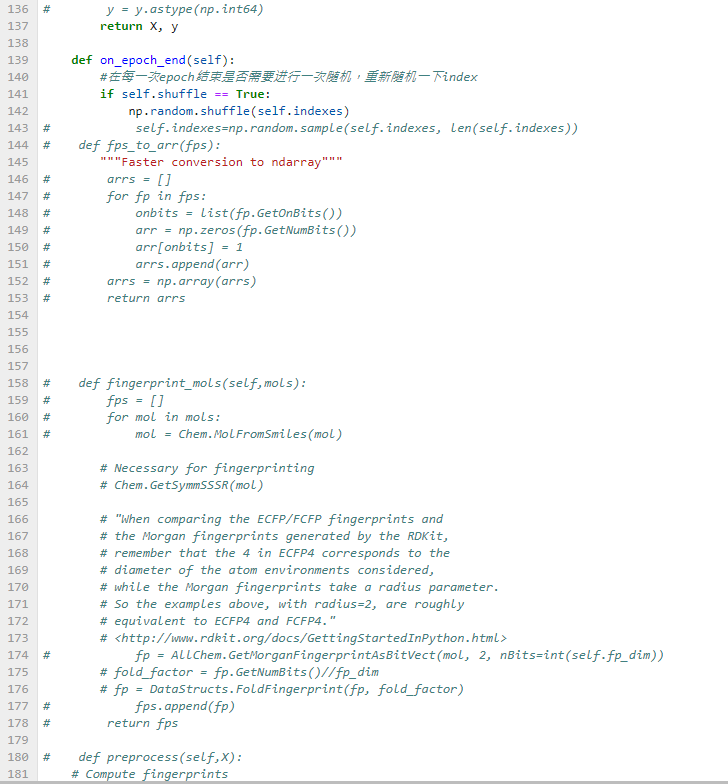
而完整程式碼請見下圖











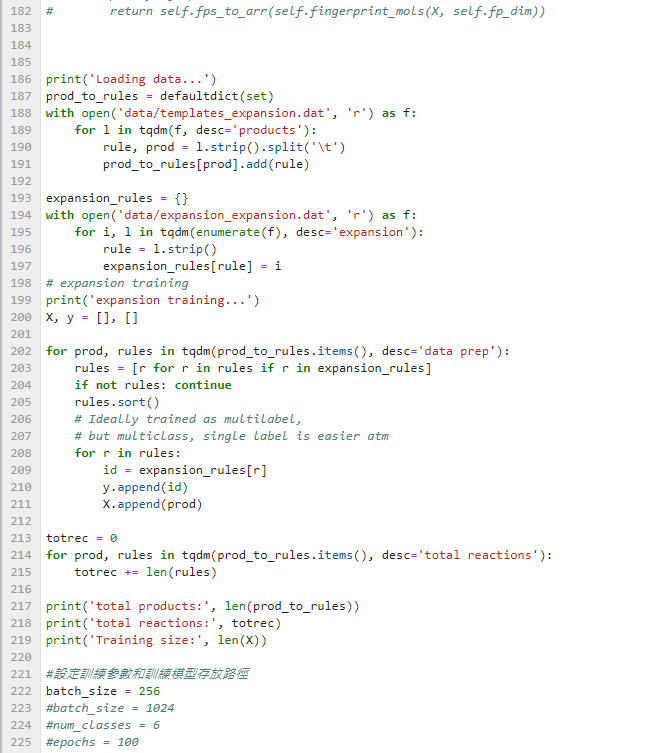








圖 9擴展神經網路訓練之程式碼

**4.2推演神經網路訓練**

簡要說明如下，一樣先用函式getmorganfingerprintasbitvect()將化學反應資料庫中各筆資料的產物分子指紋擷取出來，當作神經網路的輸入資料，但是分子指紋的向量維度比起上一節的擴展神經網路，縮小為8192，並同時拿掉五次Highway層及Dropout層連結，雖然訓練出來的神經網路效能會比較弱，但種種措施都是為了要加速神經網路的運算，之後在蒙地卡羅搜索時，才能快速的得到報酬分數，此外也改用反應的規則集Set#A當成答案的選項，並把產物分子對應的反應規則當成正確答案，來做此神經網路權重的訓練。神經網路的第一層為512個神經元的dense層，並經過ELU激活函數，而後經由Dropout層，以防神經網路過擬合，最後透過Softmax層，找出各種可能反應規則的機率分布，下圖為流程示意圖

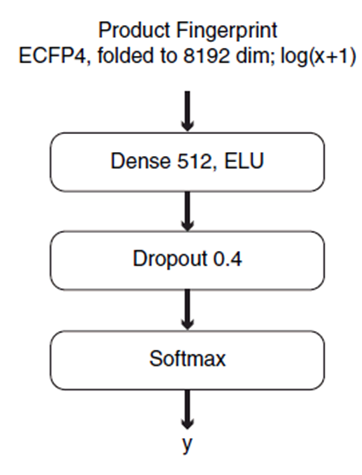
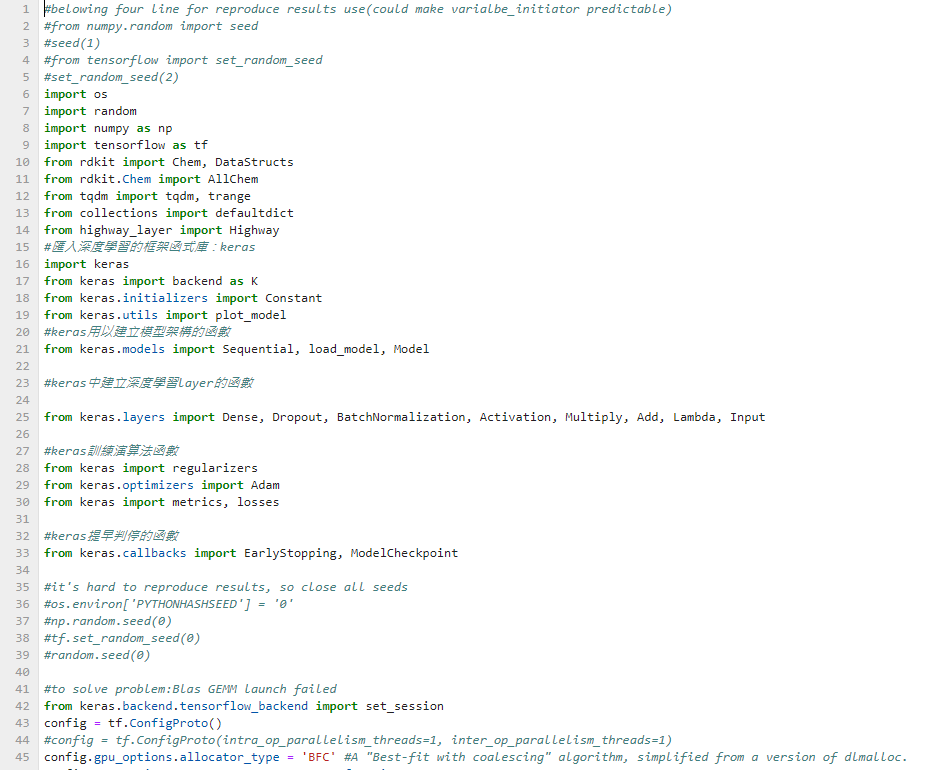
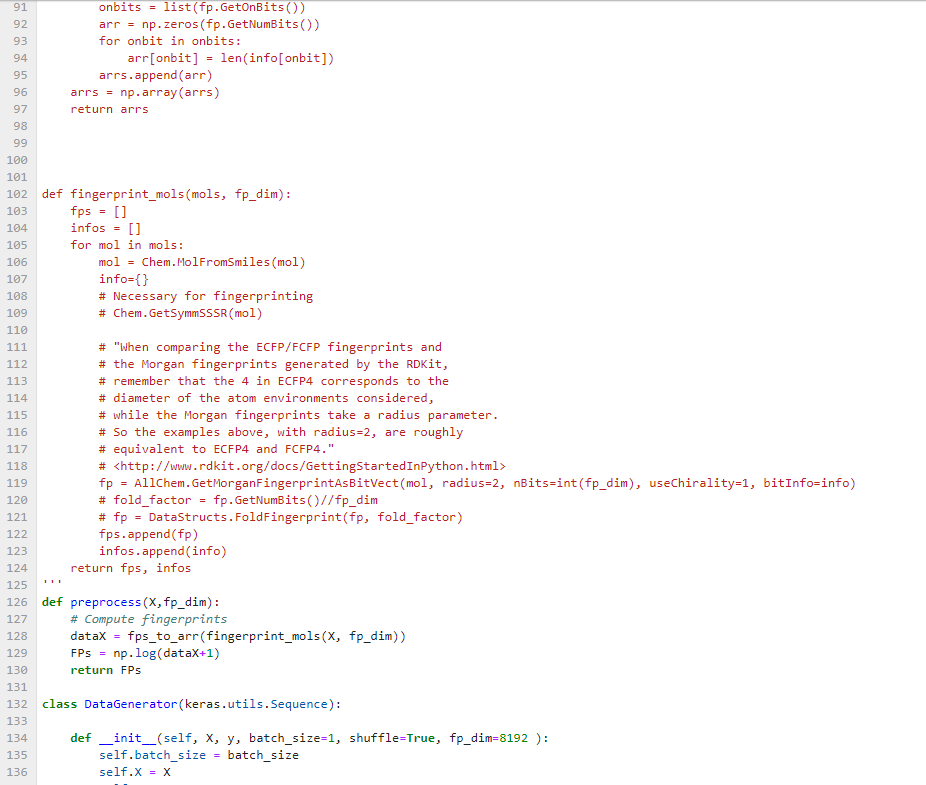


圖 10推演神經網路訓練之流程

而完整程式碼請見下圖







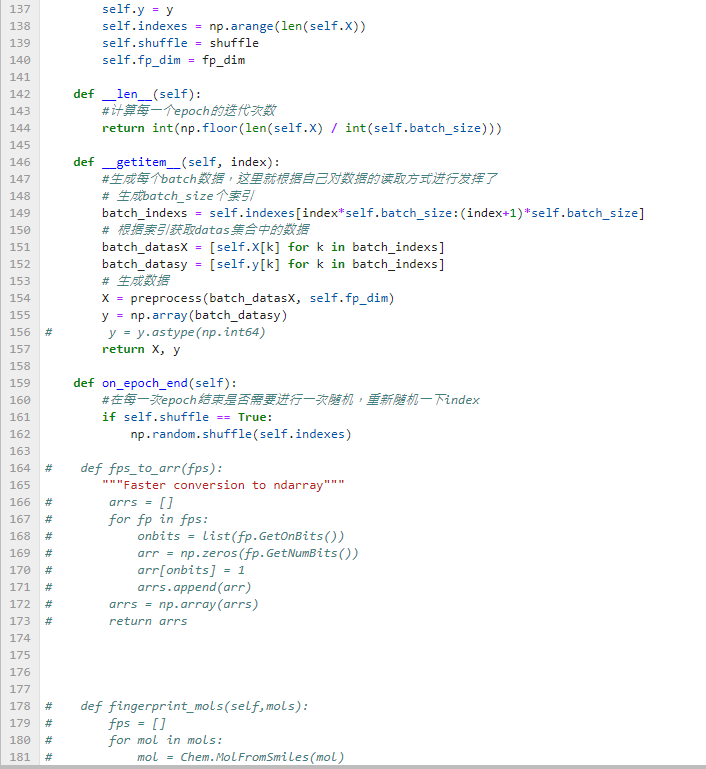










圖 11推演神經網路訓練之程式碼

**4.3 In-scope filter神經網路訓練**

簡要說明如下，先集結正確的化學反應資料庫以及負面化學反應資料庫，將正確的化學反應當成正面教材(意味著可行的機率為1)，錯誤的化學反應當成負面教材(意味著可行的機率為0)，以訓練神經網路，讓神經網路可以判斷各種產物配對特定的化學反應規則，可行的機率有多高。因此輸入的資料有兩項，分別為產物的分子指紋，以及化學反應規則指紋。化學反應規則指紋可由反應物指紋與產物指紋相減而得到，而反應可行的機率則為訓練答案。一樣先用函式getmorganfingerprintasbitvect()將運算所需的指紋擷取出來，在產物分子指紋的神經網路分支上，先連結一層為1024個神經元的dense層，並經過ELU激活函數，而後經由Dropout層，以防神經網路過擬合，而後反覆連結Highway層五次，在與化學反應規則指紋的神經網路分支做結合，而化學反應規則指紋只經過一層dense層的連結，透過cosine層與另一神經網路分支連結，最後透過激活函數Sigmoid，以判斷此產物與反應規則可配合的機率，意即反應的可行性，下圖為流程示意圖

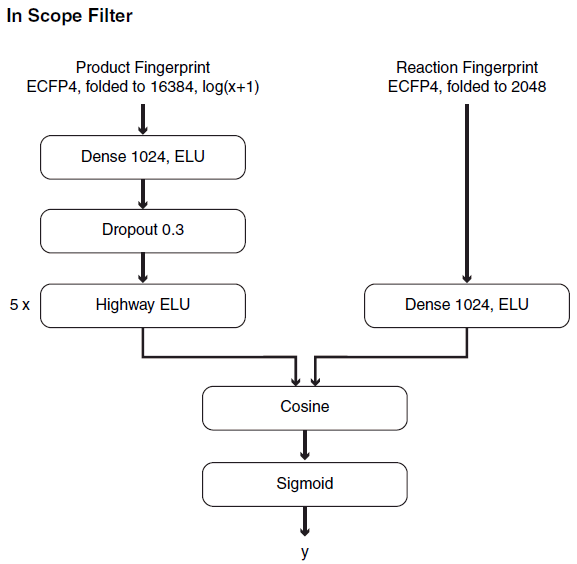
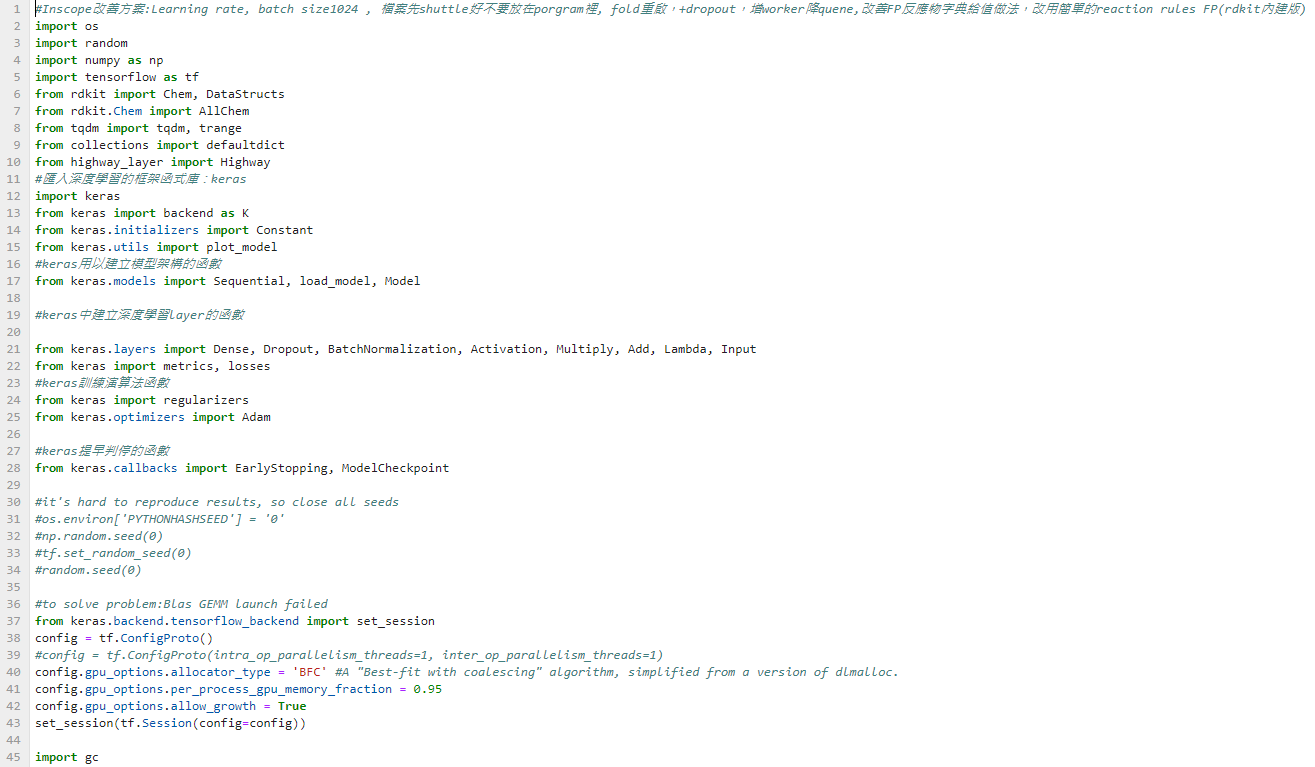
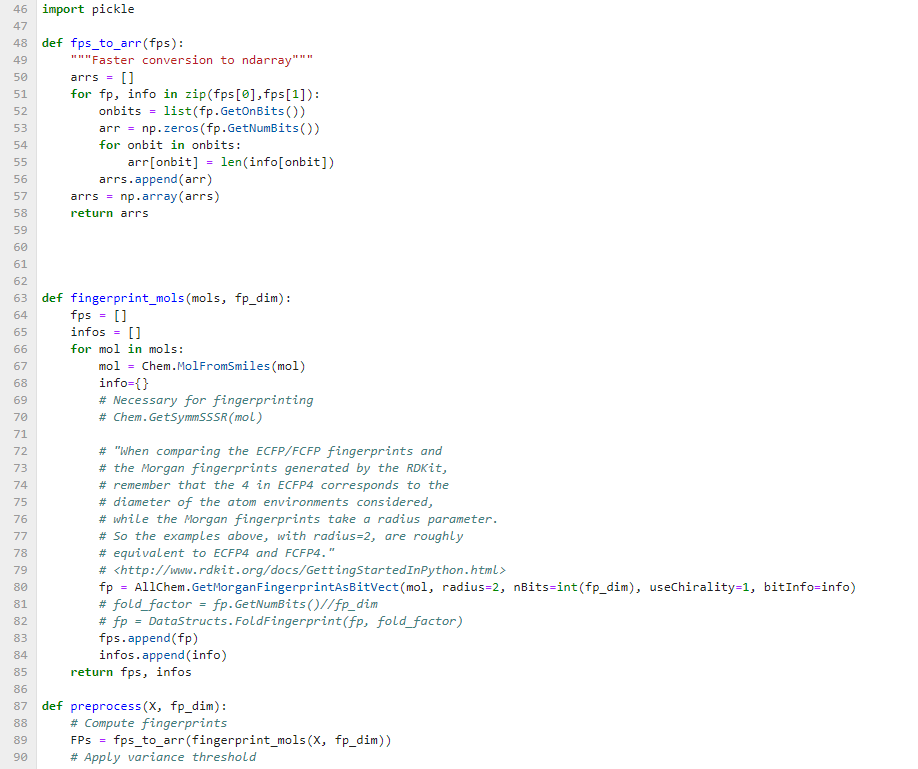


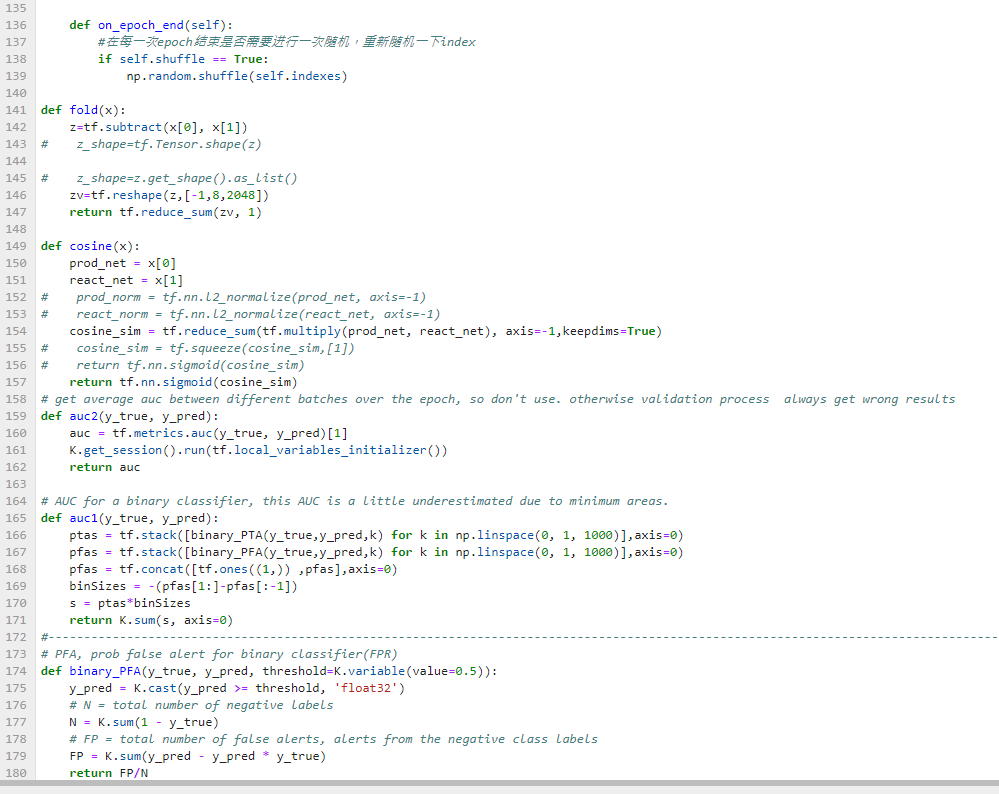
圖 12 In-scope filter神經網路訓練之流程

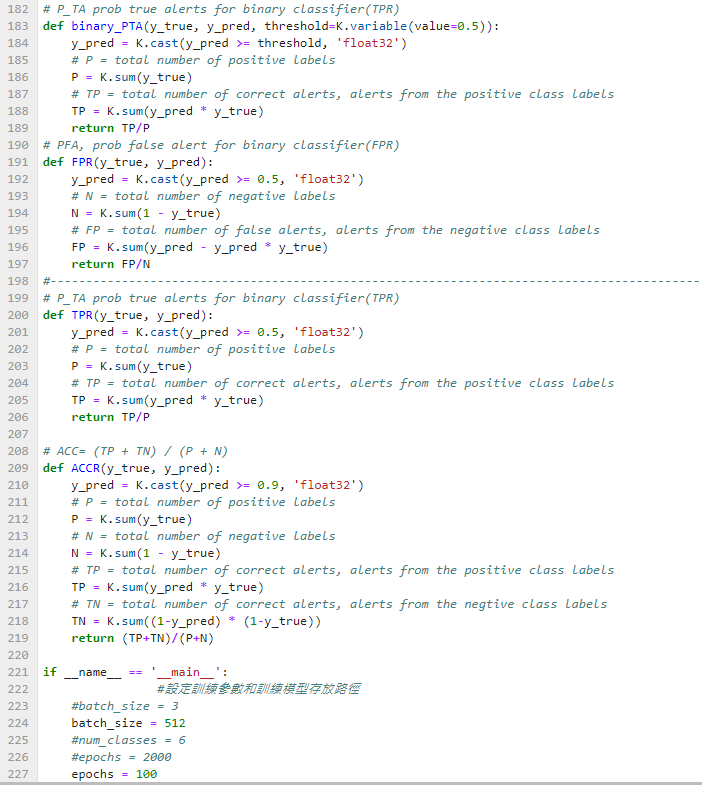
而完整程式碼請見下圖















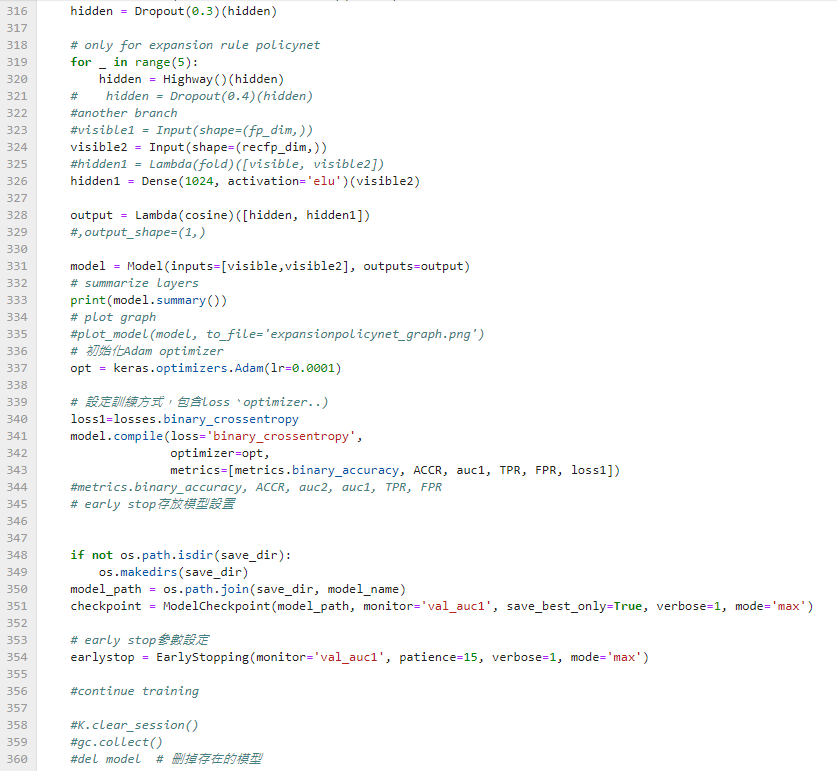
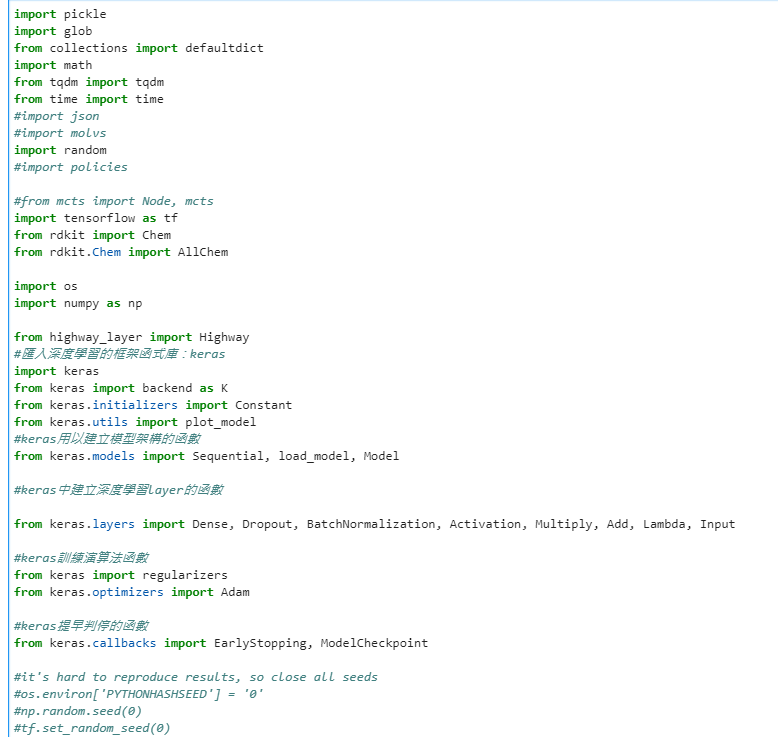


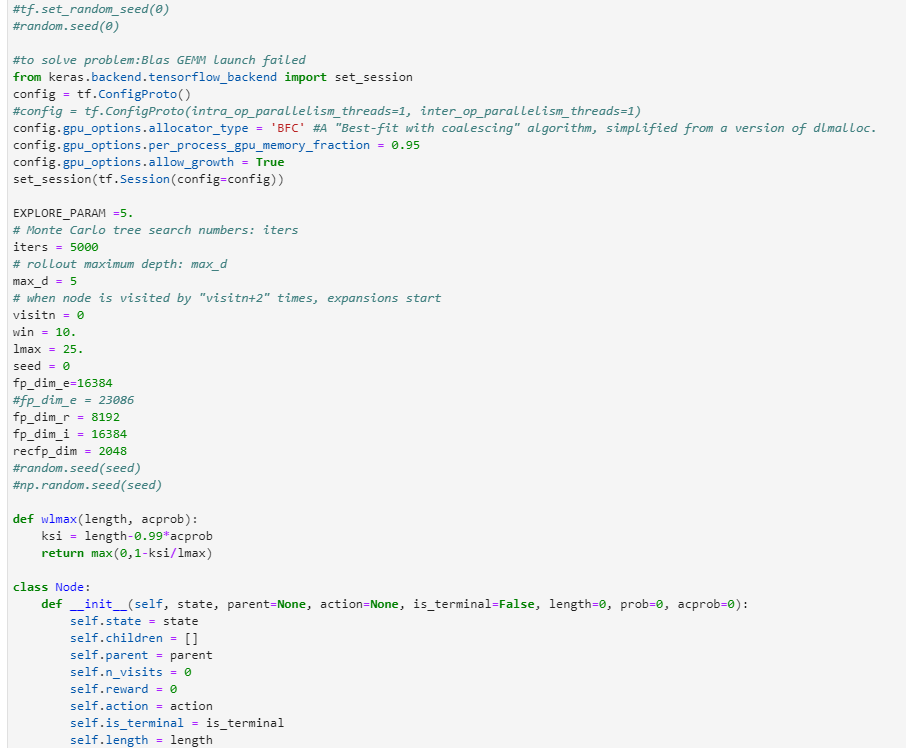


圖 13 In-scope filter神經網路訓練之程式碼

**5.蒙地卡羅搜索及逆合成路徑預測**

簡要說明如下，給定特定的化學分子，經由函式plan()進行化學逆合成途徑創建，在創建的途中利用本研究撰寫好的蒙地卡羅搜索函式mcts()針對每個節點(也就是物件Node)進行搜尋最適合的子節點、評分及記錄等行為，並逐層回傳報酬分數給根節點，反覆在搜尋最佳路徑以及探索新路徑的平衡下，最終找到合適的逆合成路徑。至於搜尋就要利用上章節訓練好權重的擴展神經網路以及In-scope filter神經網路來預測可行的反應路徑，而評分則會用到訓練好權重的推演神經網路來推估，若整個逆合成反應步驟越少，選擇某特定反應規則的機率越高，則評分會越高，完整程式碼請見下圖，



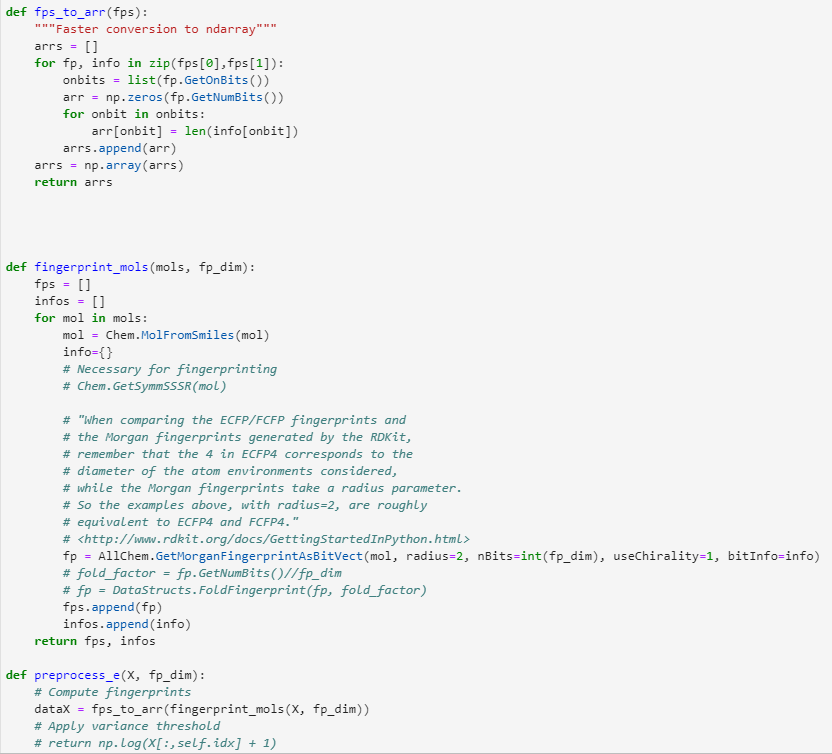












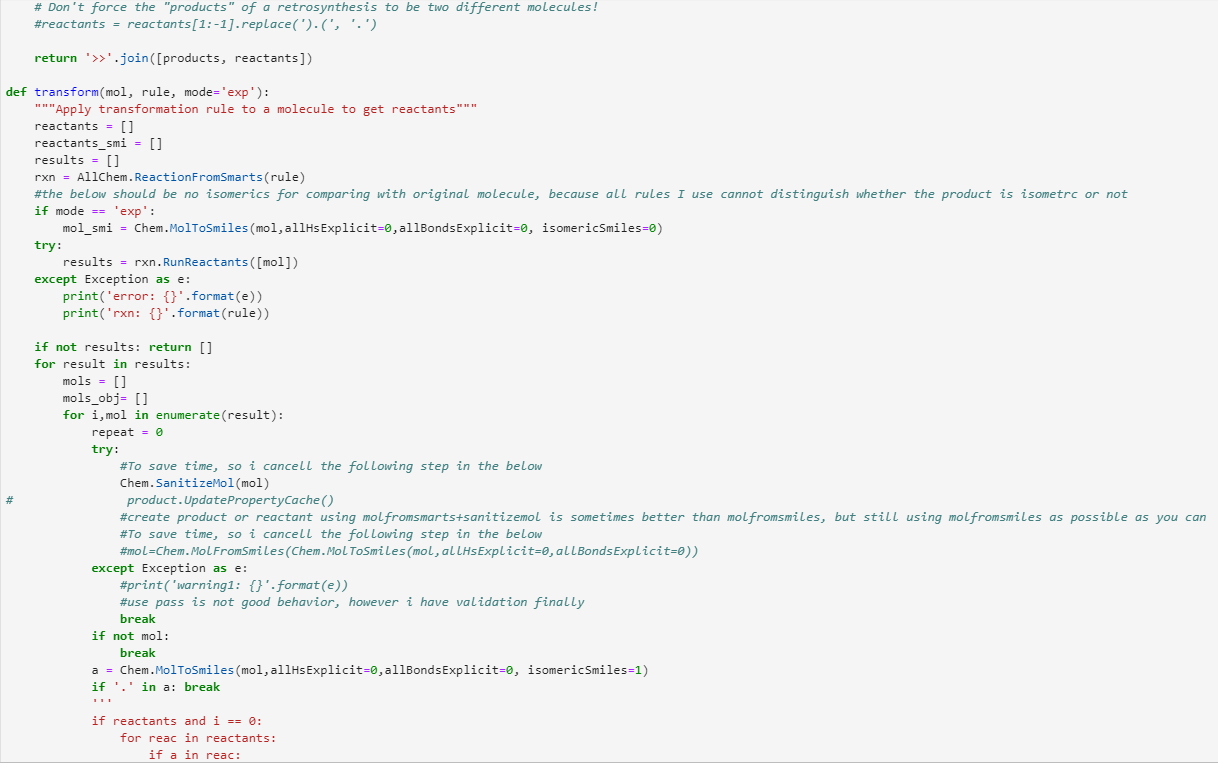




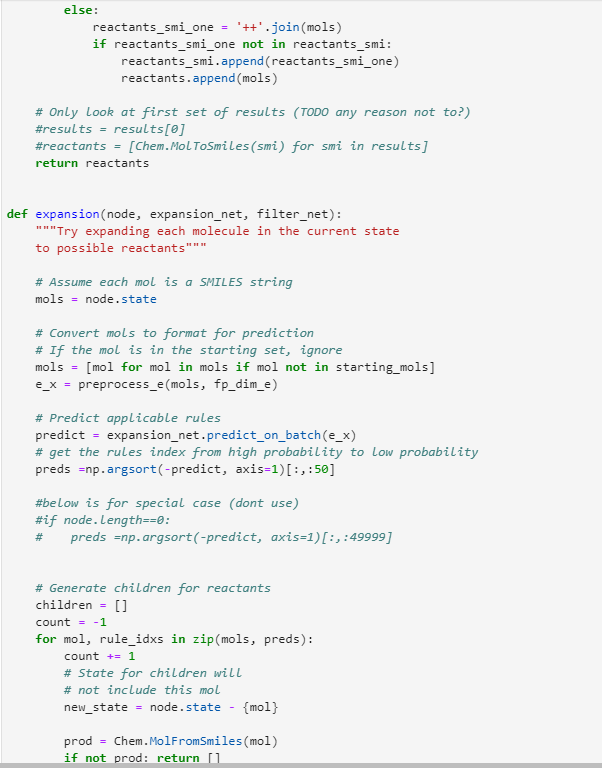




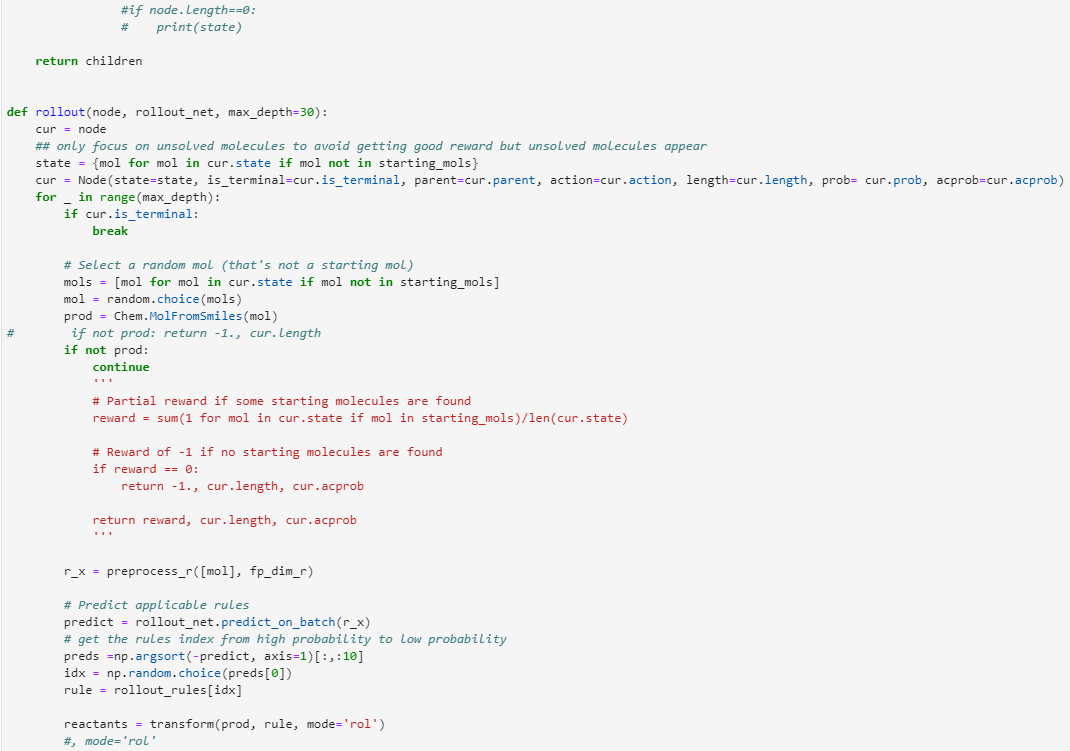


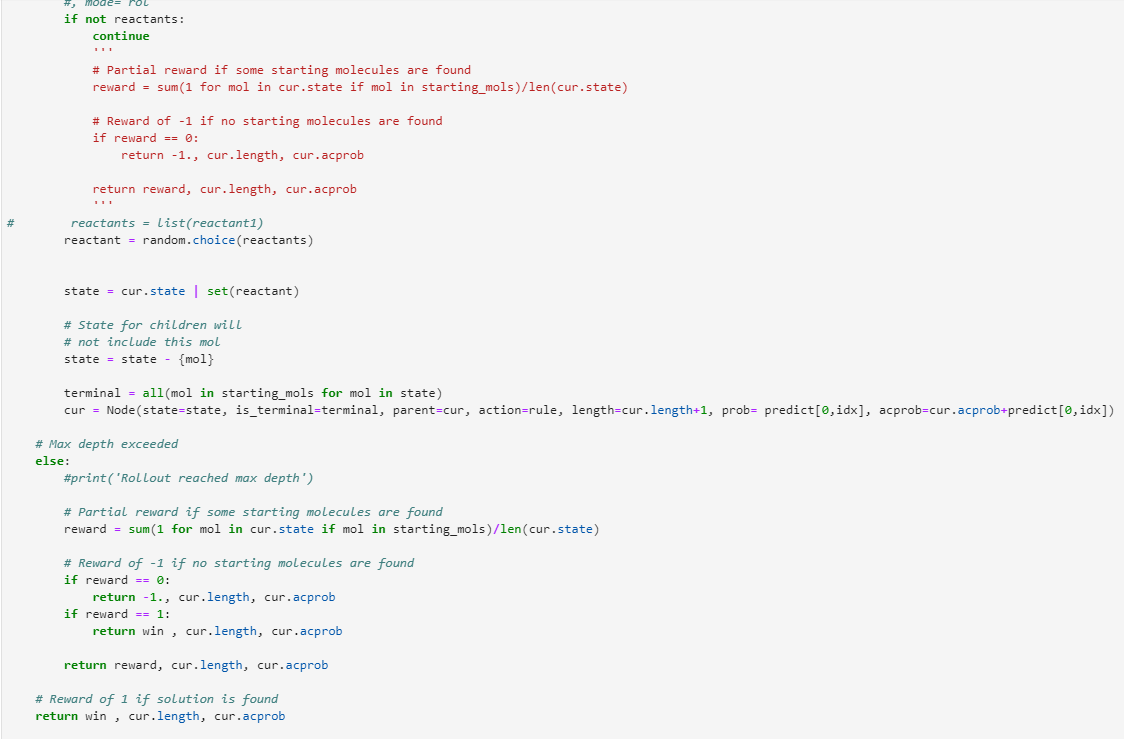


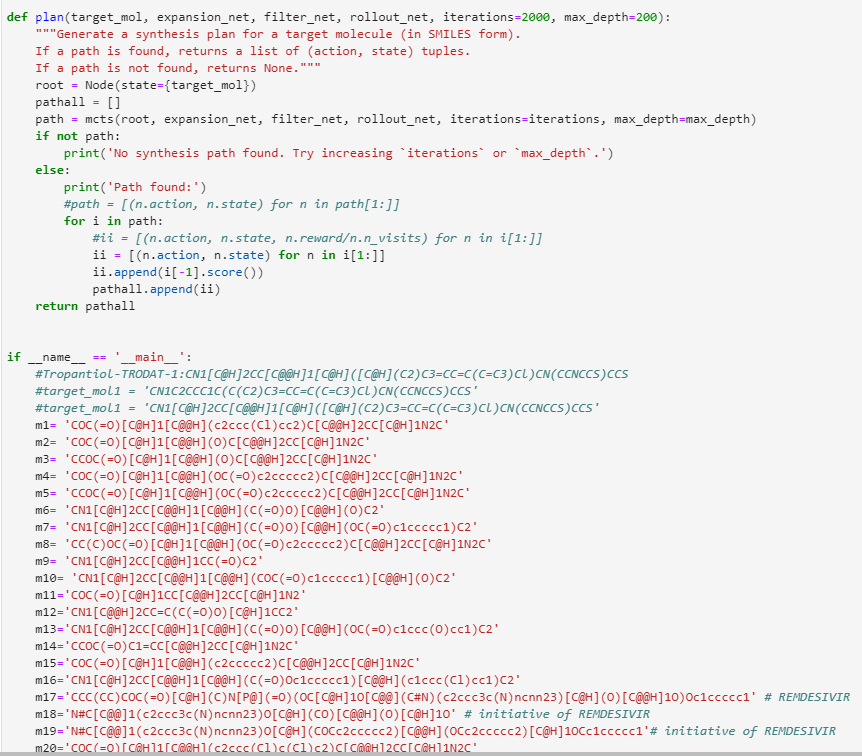


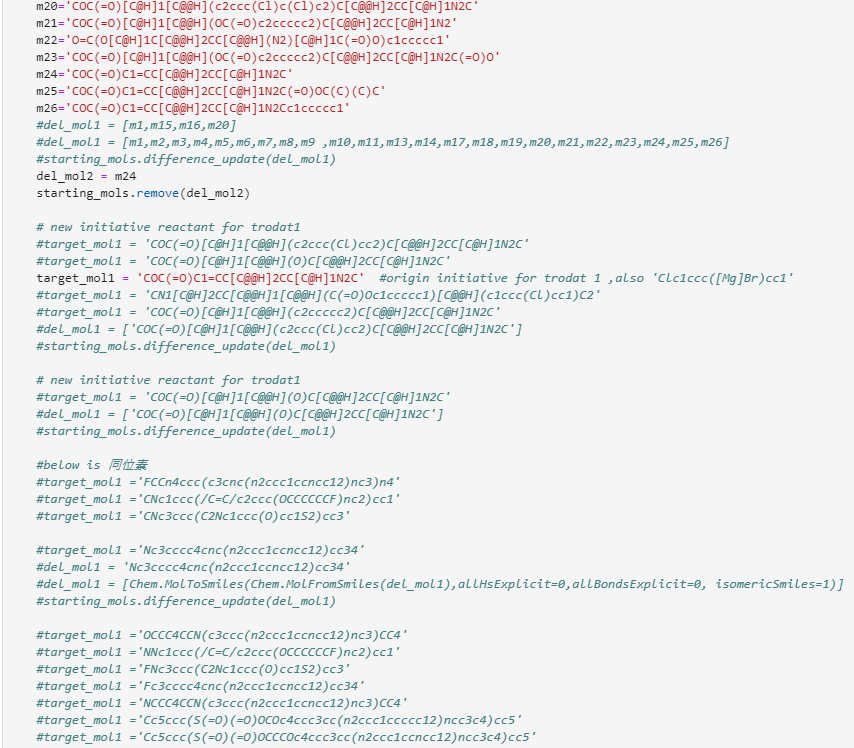












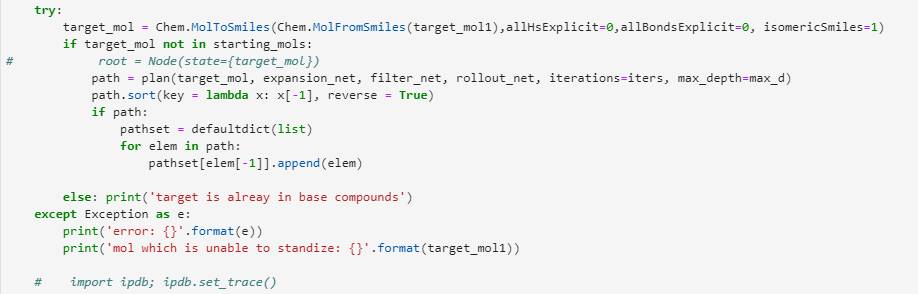


圖 14蒙地卡羅搜索結合神經網路之程式碼

**6.模擬結果(以TRODAT-1為例)與總結**

**6.1模擬結果**

以核醫藥物TRODAT-1為本研究目標測試分子，利用本研究完成的核醫藥物逆合成 AI model初版軟體進行規劃，可以產出數十條新規劃的逆合成反應路徑。經化學專家檢視，其中所規劃的一條反應路徑比商業軟體Scifinder等規劃的更具創新性，除製程步驟較核研所原製程減少4步驟外(核研所原製程為11步驟，本AI model規劃為七步驟)，規劃之起始物為非原製程之管制藥品鹽酸古柯鹼，不但可避開使用管制藥品的風險，購買及取得性也更為容易。本研究發展的AI軟體模擬規劃的TRODAT-1新逆合成及相對新製程如圖 15及圖 16所示。

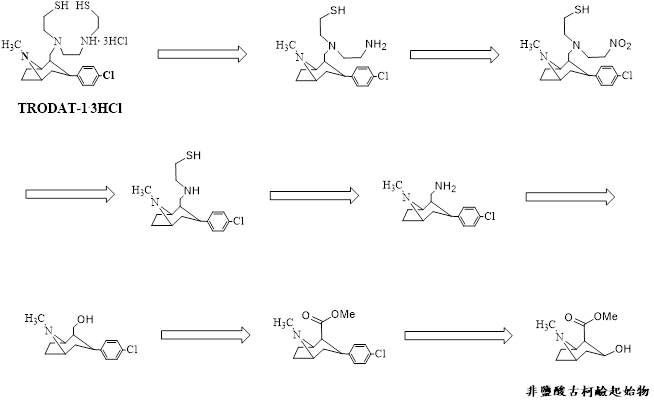


圖 15逆合成途徑預測-以核醫藥物TRODAT-1 為例

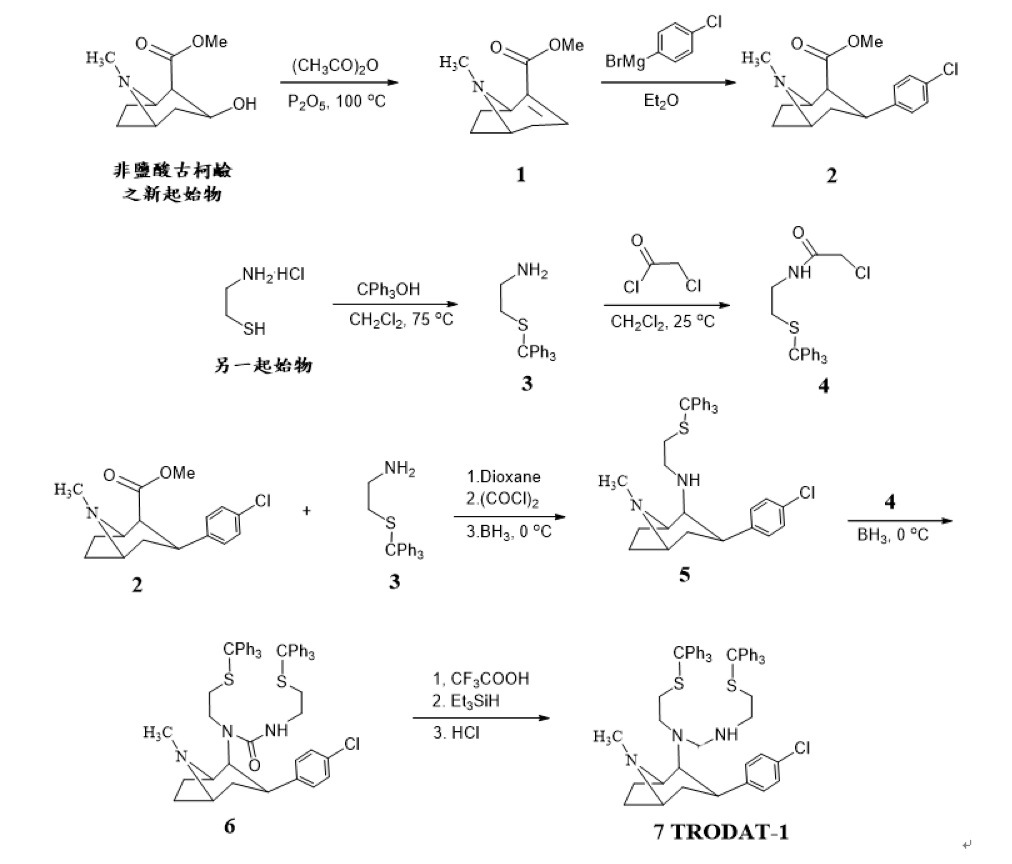


圖 16 TRODAT-1新製程規劃

有鑑於TRODAT-1原製程常有成品汞超標的困擾，本研究發展的AI軟體，除模擬規劃出前述TRODAT-1新逆合成及相對新製程外，亦針對成品汞超標問題進行探討，此問題主要源自硫原子接保護基和最後步驟去保護基時，使用了含汞試劑，所以本計畫建立的AI也對此課題，提出了無汞製程的規劃，因硫原子去保護基的反應在最後步驟，起先欲進行此項驗證時，似乎要走完整個製程，後來發現反應部分全集中在二硫二氮結構上，所以可規劃在一個含二硫二氮結構之較簡單分子上，進行AI提出的硫原子保護和去保護無汞製程之化學驗證，如下圖所示。

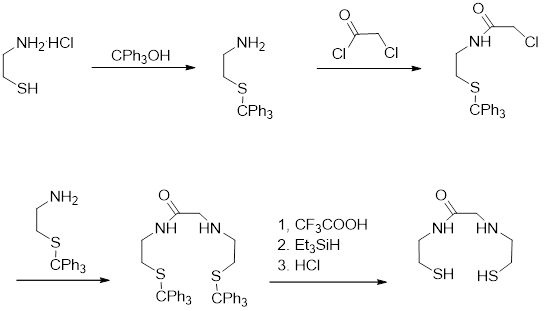


圖 17逆合成途徑預測-無汞製程建議測試步驟

**6.2總結**

本研究將Segler[[2]](#footnote-2)所提化學逆合成規畫演算法實體化，成為可執行藥物化學逆合成規劃之AI模型，利用3種神經網路經由5種化學反應數據資料庫中大筆資料的訓練，分別習得預測可能的化學合成反應及其合理性，並經由神經網路對每步化學合成反應的合理性挑選，再經過蒙地卡羅樹深度搜尋及回饋學習，看能否找得到最終的起始物基本分子，最後取得合理藥物逆合成路徑規劃。目前已完成化學逆合成 AI model簡易版軟體，後續可結合化學家專業知識及經驗，加速各種核醫藥物原料藥開發步驟更少、產率更高的新製程，並實際做化學驗證，證實模擬結果的可行性。

1. *. Med. Chem.* 2016, 59, 9, 4385–4402 Publication Date:March 30, 2016 <https://doi.org/10.1021/acs.jmedchem.6b00153> &. J Chem Inf Model. 2010 May 24;50(5):742-54. <https://doi.org/10.1021/ci100050t>. [↑](#footnote-ref-1)
2. Segler, M., Preuss, M. & Waller, M. Planning chemical syntheses with deep neural networks and symbolic AI. Nature 555, 604–610 (2018). https://doi.org/10.1038/nature25978 [↑](#footnote-ref-2)