



腦、科技與法律

——人工智慧的司法運用

課程：當代認知神經科學－腦與心智

授課老師：焦傳金 教授

學生：109091031 陳濬萱

系級：科管院學士班 24級

繳交日期：111.1.12

前言：

由於我本身的一二專長是法律與資訊工程，而且未來想著重研究的課題是「人工智慧預判」，因此本次報告主題，我將結合「腦、科技與法律」做深度的討論，並且藉由這份報告探討為什麼科技與法律的結合即便歷經多年研究，仍無法普及於日常法院判決中，希望能藉由此次報告重新審思法律（本報告將著重於刑法）建構的初衷。

與課堂的連結——精神障礙辯護：

「為什麼我們要去保護一個作惡多端的人呢？」「他們不是活該被罰嗎？」以下我將提出一些背景知識及我個人的解讀。其實，之所以會有精神障礙辯護的存在是基於刑法第19條，其立法用意是為了要確認行為人符合具有行為能力及責任能力兩項犯罪之構成要件。責任能力其實就是我們所熟知的以「年齡」作為區分依據，14以下為無責任能力，犯罪不罰；14~18歲之間或80歲以上屬限制責任能力，得減輕其刑，以及18歲以上（刑法的成年）屬完全責任能力，須完全對自己的行為負責。而在精神障礙辯護的部分比較著重於討論「行為能力」，依照刑法第19條簡單而言，行為時如果因為精神障礙或心智缺陷，致使行為人不知道其行為意義或不知道其行為將造成之結果，則理應不罰。其實這個概念再講白話一點就是刑法不會去罰一個連自己正在做什麼或自己正在做的事會造成什麼後果都不知道的人，這也是為什麼刑法不會去罰一個小朋友的原因。至於台灣為什麼會有精神障礙辯護的適用主要是因為我國立法院三讀通過了「兩公約」，並且另外制定「公民與政治權利國際公約及經濟社會文化權利國際公約施行法」，明定兩公約具有國內法的效力。而兩公約的其中一項訴求就是「限制死刑之科處與執行」，換言之就是如果想要重罰（不論是死刑還是刑期比較重的有期徒刑或無期徒刑），都必須經過嚴格、公正的審判，這樣的公約其實就是變相的在打壓死刑，以達到保障人民生存權利的目的。

「既然加害人也是人，那麼也同樣的，我們也必須給他保障。」看到這邊，相信大多數的民眾又會開始忿忿不平，「殺人償命，天經地義」等等的思維漸漸浮出檯面，而這個想法或許沒錯，也應驗著由應報理論¹延伸出的罪責原則²概念，不過，大家或許忘了他們口中的恐龍法官最需要顧及的其實是所謂的「公平性」。長年以來，社會大眾不斷將做錯事所以要受到處罰的想法固化，再加上新聞媒體的渲染，人們開始將刑罰的存在視作是受害者用來追訴犯罪的一種工具，但是事實並非如此。刑罰的存在是為了代表國家去罰那些違反國家規定的人，國家基於保障社會安定性，在人權與公義之間衡平，最終選擇犧牲部分的人民權利，來換得整體性的公益。當然，如果換個角度來看，刑法表面上雖然是限制了人民，但大眾卻也因此得以避開弱肉強食、自然法則下的世界，也算是藉由這樣的限制讓多數人能享有更多的權利。當個體破壞了國家刻意製造的衡平結果，使國家對你期待落空，而在國家不能選擇不保障誰的情況下，國家只能用懲罰的方式來指責你破壞了他對你的信任。綜上所述，刑法存在的概念其實就是透過反面解釋的方式去告訴人們你們可以做些什麼，而刑罰的存在則是為了幫國家罰不聽話的人，而不是為了幫被害人罰欺負他們的人，因此我們並不能要求刑法完全站在被害人的角度去想，因為被害人的心態並不是刑法在意的點。上述兩者的主體都是國家對人民這種上對下的權利關係，而在雙方地位不平等的情況下，國家當然就必須給出相對應的優待，以達成實質平等，這正是為什麼會有精神障礙辯護的出現。

「精神障礙辯護會不會成為加害人的擋箭牌呢？」我認為是不會的。綜觀台灣史上所有提過精神障礙辯護的案件，可以發現大多只會出現在被判「重刑」的情況下，換言之，他是辯護者為求壓低刑期的最後手段。其實就像老師上課提到的，精神障礙辯護必須經由一定的程序提起，且提起後須經專家多重鑑定，並不是申請了就一定會過。再者，判定一個人有沒有精神病之結果不會改變判刑標準，標準就是那樣，做出多糟的行為，就要為自己的行為負多大的責任。精神障礙辯護只不過是國

¹ 應報理論(Retributive Justice)認為因果報應是自然的理性，而刑罰是犯罪的存在即是一種報應。

² 罪責理論指的是一罪一罰，簡而言之就是人要為自己的某個行為負起某個責任。

家會了平衡自己與人民之間的地位而做出的些微妥協罷了。大眾對於精神障礙辯護的反彈很可能並不是來自於制度的本身，而是因為媒體不斷地去加重討論這項議題，使得潛在的犯人認為「精神障礙等於不能判死刑」，導致只要是重大刑事案件，很多犯人都會刻意去模仿、去濫用這樣的辯護手段。在濫用與媒體的渲染下，大眾只會更加反感進而誤解此項辯護行為的真正用意，我想這也是法律人想靈活運用科技所需克服的難題吧！

在判決中使用科技³：

課堂中，老師介紹了fMRI於法律判決中的應用，而由於說謊成性的人不論是在說謊話或說實話，其前額葉認知功能區皆會活化，因此帶出了fMRI無法作為測謊工具的結論。除了測謊之外，實務上也有使用過COMPAS統計模型⁴來判斷被告的再犯率，如果結果是高程度的再犯率，那麼很可能會從重量刑。COMPAS的統計模型雖然是比較簡易的演算法，不是那種很精細的人工智慧，但於我而言已經算是一個引入科技進入司法程序的先驅，他最早是2016在美國先起用的，目的就是會了藉由自動化提升司法效率，值得注意的是台灣的人工智慧判讀也是在2016年做左右開始慢慢有越來越多人研究的，可見COMPAS的試用成功對於人工智慧引入司法程序有相當重要的影響力！（好振奮）那麼究竟什麼是COMPAS演算法呢？它的概念其實很簡單，用等式表示大概是：

$$s = a(-w) + a_{\text{first}}(-w) + h_{\text{violence}}w + v_{\text{edu}}w + h_{\text{nc}}w$$

其中， s 就是我們要求的再犯率， a 則是犯人現在的年紀， w 是一個權重因子， a_{first} 是第一次被逮捕的年紀， h_{violence} 是施行暴力行為的歷史紀錄， v_{edu} 是受教育的程度，最後的 h_{nc} 是不遵守法律的歷史紀錄。需要特別注意的是，上述的數值都是採用全國性的數值，並非個案，也就是說算出來的再犯率會是所有類似的人的再犯率，並不是專屬個案當事人本人的再犯率。

從上面的等式可以看出他其實就是在用一些歷史資料來進行預測，不過這樣的演算法適用卻引發了軒然大波……而這就是我想去討論的「為什麼人工智慧在司法中的運用永遠無法有所突破」的議題。首先，大家爭論的點普遍在於「違反正當的法律程序」，憲法中明文規定量刑必須是基於「正確的資訊」，但是若是採用COMPAS模型，因為數據全部都是交給電腦去跑，其實很難確定到底他的運作方法會不會是正確的，而且事後也不能用交叉詰問的方式去比對正不正確，這樣將會影響到日後的救濟程序。另一個爭論點也是在於**量刑理應個人化**，運用這個數值算出來的是「平均而言」，也就是說他已經預設了只要有這樣的背景，就一定會是這樣的再犯率，他並沒有考慮到個別的差異，只是將類似的人全部劃分在同一個區塊而已。另外還有，大眾普遍猜測程式的設定很可能會**不夠公正**，其一是因為很難保證設定程式的人沒有個人喜惡或是收受賄賂，而確實，實務上就有研究指出非裔人士被標註成高再犯率的機率是高加索裔的兩倍；其二是因為當初在設計COMPAS模型的時候，就有專門為女性設計一個量表，而大眾普遍認為這樣將獨利於單一性別，並**不符合性別平等權**。最後就是在涵攝案件時常會有各種不同的「劇情」，而當運算過程**不夠透明化**時，人們很難確認是否真的有考量到完整事件經過，還是只是操縱者以自己個人的認知去操控運算而已。我想上述這些問題都是普遍大眾對於科技應用的不信任點，其實不只限於COMPAS，很多人工智能預判的機制都是因為上述的問題而無法更進一步去施行，我想我也可以理解民眾擔憂的原因，不過我想提出一些個人意見供與大家思考。

首先是害怕資訊不正確的部分。這部分我想採用State v. Loomis一案中法官的見解，該名法官認為所有用來跑演算法的數據都來自於被告本人提供以及公開的資料，因此正確性是可以被反覆驗證的（數入資料，跑出結果），是以相比法官自己吸收理解正反方的辯論，藉由模型跑出的數據錯誤率

³ 可以用來反思的資料：<http://www.hss.ntu.edu.tw/upload/file/201910/b4119d7a-6850-4c10-8a47-bbd4df7e73fe.pdf>

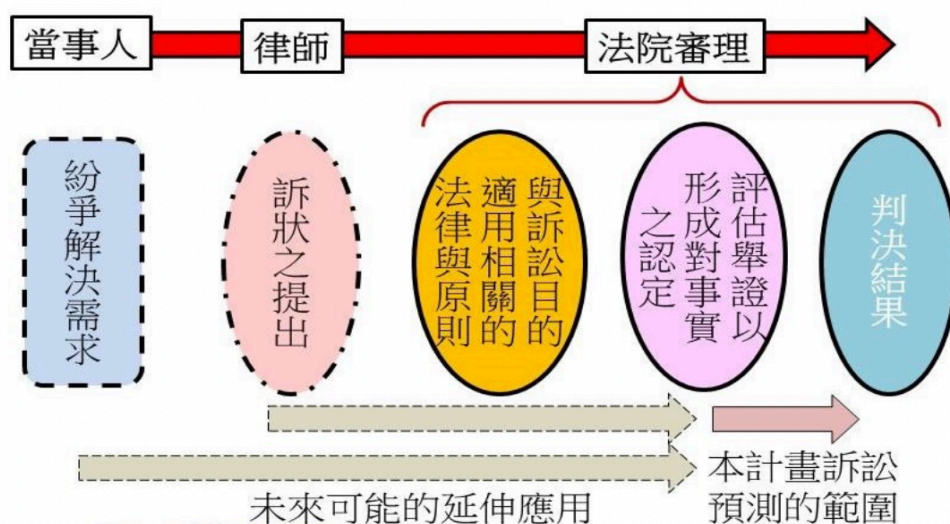
⁴ 取自：[https://en.wikipedia.org/wiki/COMPAS_\(software\)#CITEREFNorthpointe2015](https://en.wikipedia.org/wiki/COMPAS_(software)#CITEREFNorthpointe2015)

應該會更低，畢竟不會有腦筋轉不過來，或是法官當天過度疲累等個人因素導致的邏輯跳躍。對於資訊不正確的疑慮，於我而言比較像是判決確認後來自被告的不服氣（因為自己被判重刑，所以就說模型不對），換言之，如果說今天被告被判定再犯率極低，應予輕刑，則被告很可能不會提出模型有資訊不正確的疑慮。再者，關於「量刑應該個人化」的部分，我認為適用此演算法的真正用意是用來增加司法效率，而不是全權取代司法人員。也就是說，使用COMPAS模型的同時，法官基於案件獨特性，仍會依照個案去做更深入的判斷，因為對於法官而言模型的適用只是一個可以參考的數據資料而已，真正的判決結果仍然是由法官自由心證。而關於程式可能不公正的問題，因為自己並未查到整個程序的誕生形式，所以不能確定程式的撰寫是否真為公正，但是我認為如果有這方面的疑慮，其實可以考慮在司法院特別開設一個部門，由司法人員親自去學習如何跑數據，如此一來或許能大幅提高人民對於該程式的信賴度。另外，也想請大家思考看看「難道全由法官去判斷就一定是客觀公正的嗎？」從這個問題其實也可以帶出「透明化」的問題，我認為不論是程式或由法官判定基本上都有一定的不透明性，但若真的要進行比較，使用COMPAS模型好像更透明化一點。原因很簡單，法官是人，人都有自己的思維模式以及不同的認知，因此在面對相似的問題時，每一位法官所作出的判決都可能有所出入，這也正是為什麼救濟程序中准許當事人能夠上訴並且在窮盡一切方法後能申請大法官釋憲。COMPAS模型不同，資料（於庭內）是公開的，演算公式也是公開的，特定input會導出特定output，因此不會有自由心證的問題。

台灣如何將人工智慧運用於司法實務：

老師在課堂中提到的案件以刑事案件為主，但事實上，台灣在人工智慧預判的部分並不普及（我們從德國繼承了十分傳統的歐陸法><）再加上台灣人對於科技於司法中的運用仍有所質疑，因此目前實務上，科技只被用於較不會侵害人身重大權利的案件，大多是民事案件，尤其是親權訴訟。在這邊我想分享目前清大科法所的 Civil Law Project⁵來做為這份報告的結尾！這個計畫搜集了2620篇2015-2017年間各級法院所有親權案件，包含離婚、子女最佳權利等判決，並將這些判決資料做成等效資料，之後再去用自然語言處理來預測判決的研究。（下圖取自科法所AIFR GROUP）

民事判決書標註與「訴訟預測」的關係：



簡單介紹一下，這個研究主要是使用結巴(jieba)中文分詞器外加中文法律辭典來進行中文的斷詞，而語意分析的部分則是使用非監督式的隱含狄利克雷分布法(LDA)以及類神經網路的自動編碼(auto-encoding)技術。等到拿到案件的編碼後，會透過深度神經網路(Deep Neural Network)等機器學習方法來進行文本的判讀、分類與預測。藉由上面的程序，我們可以把某一類型的判決書與判決結果

⁵ 取自：<http://www.phys.nthu.edu.tw/~aicmt/Civil%20Law%20Project.html>

當作訓練資料，讓機器學習從文本的詞彙與主題來進行預判。這種方法我覺得能達到跟上面的COMPAS蠻相近的結果，也就是說兩者都是一個不倚賴人類主觀參與的、具有獨立性的研究方法，如果能達到一定的準確度，想必就可以回推法官在判決書中的用詞與主題間的關聯性，還可能可以用來當作判決品質的判斷依據！是不是很振奮人心！

反饋：

雖然礙於篇幅過長的問題，我只能粗略講一些我的想法以及介紹Civil Law Project，但最後我還是想提一下大多數法律人最擔心的問題——「司法人員會不會因為人工智慧的運用而失業啊？」其實很多新聞報導中反彈聲量多是來自於怕自己失業（畢竟台灣法律系畢業的學生真的滿街都是……），但這邊我只想說，**請相信自己的專業！**機器的魅力在於準確性以及效率，人腦的魅力在於我們同時具備了理性與感性。在腦與心智的交互作用下，我們懂的什麼叫做個案獨立性，我們知道每個案件當事人的背景都有所不同，因此若只是採用統一標準，將不利於特定當事人。舉個例子，如果機器全面取代人類，則母親殺嬰屬於殺人罪，沒有過多轉圜的餘地。但是法官（人類）卻能夠考量到當時農村背景下婦女被要求生子以充當田中勞力的壓力，可以去運用自己的同理心換位思考當一位婦女看到懷胎十月生出來的是女兒而家庭經濟負擔無法負荷的崩潰情狀下，只能忍痛掐死孩子的無奈。因為由人類作為司法人員，所以有了母親殺嬰罪減輕其刑的誕生，也才使的法秩序不過度壓抑人民，這是人類在司法中的不可替代性，也是我們該引以為傲的地方。因此，人工智慧的運用從來不是為了取代人類（我是指法律層面啦……其他行業我無權置喙），而是為了讓整個程序更加有效率、提高準確性、提升判決結果的品質以及可預測性，我甚至在想，如果能普及人工智慧預判的概念，或許根本就不需要考慮陪審團的存在必要性了！因為一般民眾也都會懂得什麼樣的條件下會發生什麼結果，那麼未來將會是由一般大眾一同監督司法的運作，如此一來能更加確保人民權利受到保障。我想，大家不應該妄自菲薄的去排斥一個新型科技成果的運用，與其去擔心，不如去想想如何運用才能達到最大效益，也不要因為自己的猜忌浪費了這麼的研究成果！相信在人類與科技的互助下，未來的社會必定會更勝於今！⁶

以上就是我的期末報告～非常感謝您的耐心閱讀，預祝您有個健康且充實的寒假！

⁶ 可以讀看看的補充資料：<https://ai.iias.sinica.edu.tw/ai-child-custody-decision-prediction/>