資料來源

實驗資料是來自於加利福尼亞大學爾灣分校(University of California-Irvine)所建構的UCI資料庫系統中心。這個實驗為帕金森氏症患者的行走實驗，實驗室中會記錄病人的步態資料，分別要求好發凝凍步態的受試者在各個部位穿戴內建微型電腦的加速度感測儀（綁在脛骨、大腿以及臀部）後，沿著研究路徑進行十分鐘實驗，其中動作包括隨意走動、轉彎和停止，並記錄 x、y和z方向的加速度。過程中，凝凍步態的辨別是依照身旁物理治療師的觀察。在實驗過程中，加速度感測器會以64Hz的頻率紀錄資料。

資料處理

資料平衡

資料分類分成三類，分別為不屬於實驗範圍(資料中以 0 來表示)、沒有發生凝凍步態(資料中以 1 來表示)、發生凝凍步態(資料中以 2 來表示)。

由於原始資料中有大量的 0，若將所有 0 都放進模型裡，分類器會頃向猜測 0，使得特異性會過配適。除此之外，因為在資料的頭部跟尾部是接近實驗開始跟結束時，此時的資料大部分都是 0，我們認為這些資料不是我們想要分析的資料，因此將資料去頭去尾。去頭的定義為: 從頭開始算，直到遇到第一筆 1 時，在該筆 1 前面的 0 全部移除去尾的定義為: 從最後一筆資料開始算，直到遇到第一筆 1 時，在該筆 1 後面的 0 全部移除

視窗切割

收集完的資料為單一時點的加速度資料，而為了觀察出較完整的現象，我們將多個資料點整合為一個視窗，採用的是滑動視窗法(Sliding Window)。在Marc Bachlin 論文中，將資料的視窗定義成 4 秒一個視窗，且視窗間的重疊長度為 3.5 秒，也就是說視窗一次滑動 0.5 秒。但為了增加資料量，我們縮短了視窗的時間，將視窗長度調整為 1 秒一個視窗，並且視窗間的重疊長度調整為 0.75 秒，也就是說視窗一次滑動為 0.25 秒。

視窗定義

在參考過往文獻之後，我們總共將視窗分成 4 類。首先我們先將資料點以片段法來進行分類，資料分類為 2 的部分我們便將其歸類在凝凍步態片段(FOG segment)，而在凝凍步態片段的前兩秒我們將其歸類在前凝凍步態片段(Pre-FOG segment)，而剩下的部分我們便將其歸類在非凝凍步態片段(Non-FOG segment)。

一張含有 文字, 天線 的圖片

自動產生的描述

分類示意圖

接著我們再依據片段法的歸類結果來定義視窗，總共會分成 4 種情況

1. 非凝凍步態(Non-FOG): 只要視窗有包含到非凝凍步態片段便將該視窗定義成非凝凍步態
2. 前凝凍步態(Pre-FOG): 若視窗全部皆落在前凝凍步態片段，便將該視窗定義成前凝凍步態
3. 前凝凍步態過渡期(Pre-FOG Transition): 若視窗前段落在前凝凍步態片段，而後段落在凝凍步態片段，則將該視窗定義為前凝凍步態過渡期
4. 凝凍步態(FOG): 若視窗全部皆落在凝凍步態片段，則將該視窗定義為凝凍步態。

定義目標分類

針對不同的目的以及結合過往的文獻，我們將預測目標分成 6 類

一張含有 桌 的圖片

自動產生的描述

Case1將 FOG 發生前的部分也都完整地列入，認為 FOG 發生前可能也存在一定的跡象，將其加入後或許能可提升分類準確度。Case2 僅考慮 Pre-FOG，模型的主要目的是辨識出 Pre-FOG，讓機器能夠最即時的預測將來可能發生的危險，起到保護患者的功用。Case3和case4也是以在 FOG 發生前就能預測為想法，希望將Pre-FOG結合Pre-FOG transition 之後能提高準確率。Case5則是希望在患者發生凝凍步態的當下能夠有效辨別。而Case6則希望將Pre-FOG transition包含進去之後能夠提升準確率。

我們將研究重點放在Case3與Case6，除了因為這兩個Case相對較有解釋力以外，也可以讓我們同時了解凝凍步態發生前以及凝凍步態發生當下的情況。

資料再平衡

在定義完視窗之後，我們發現資料不平衡的情形還是相當嚴重，為了便免分類器產生偏誤，因此我們針對切割完的視窗進行資料再平衡。對於每個case我們皆採用相同的平衡方法。首先我們計算 Target Class 的數量，接著從 Nontarget Class 中抽出 Target Class 兩倍的視窗數。

特徵擷取

我們所使用的特徵可以分為三大類。

1. 根據原始加速度值得到的敘述性統計

一張含有 文字 的圖片

自動產生的描述

2. 原始資料經過傅立葉轉換後得到的敘述性統計及能量

一張含有 桌 的圖片

自動產生的描述

3. 配飾樣條函數所得的係數