

機器學習模型設計過程和MEMS MLC

作者：意法半導體

開發機器學習專案的五個步驟 — 掌握要點，應用並不困難！

邊緣機器學習具有許多優勢。然而，由於開發方法與標準程式設計方法截然不同，許多機器學習開發者可能會擔心自己難以駕馭。

其實，完全沒有必要擔心。一旦熟悉了步驟並掌握了機器學習專案的要點，就能夠開發具有價值的機器學習應用。此外，意法半導體（STMicroelectronics，簡稱ST）提供解決方案，以促進邊緣機器學習得到廣泛應用發揮全部潛力。本文描述了機器學習專案的必要開發步驟，並介紹了ST MEMS感測器內嵌機器學習核心（MLC）的優勢。



嵌入式ST感測器的機器學習核心

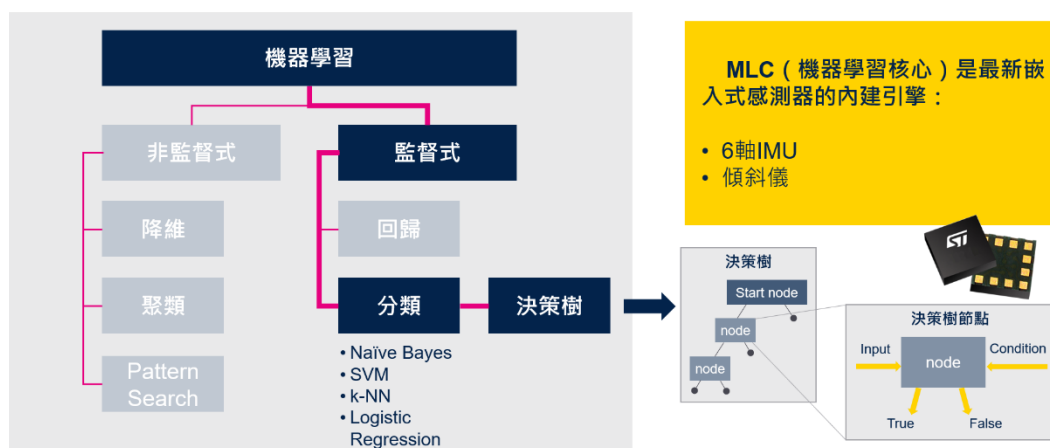


圖1、嵌入式ST感測器的機器學習核心

首先，針對機器學習模型--決策樹（Decision Tree），該模型與MLC一起內嵌在ST MEMS中。

當使用者想要在嵌入式系統中使用機器學習或深度學習功能時，通常可按照以下五個步驟進行開發（圖2）。

AI專案開發的5個步驟

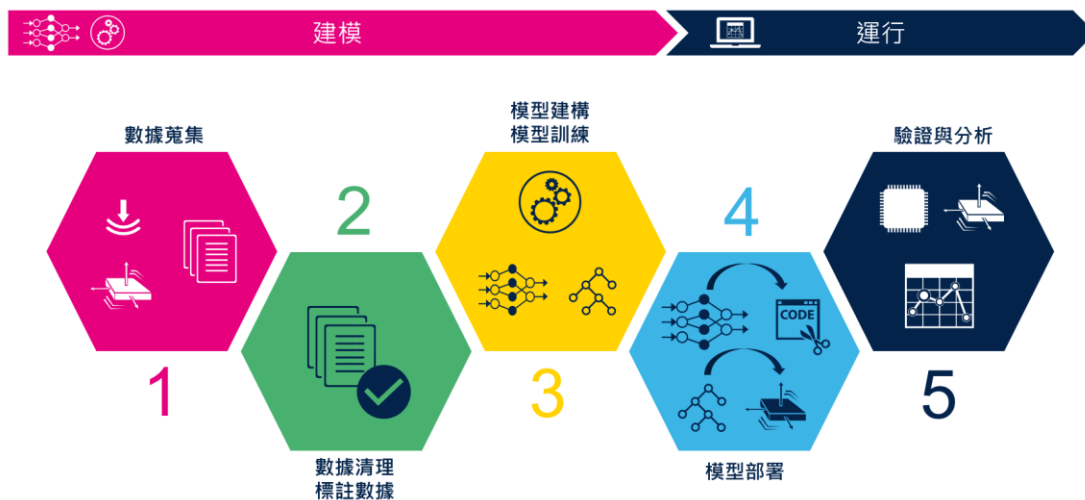


圖 2、AI專案開發的五個步驟

在步驟1中，開發者需要收集機器學習所需的數據。該數據集的一部分（訓練集）將用於訓練模型，另一部分（測試集）將用於評估所構建模型的性能。機器學習的數據集中典型的拆分比率為：訓練集80%，測試集20%。

在步驟2中，需要對數據進行標註。開發者需要將收集到的數據標註類別（例如「跑步」、「散步」、「靜止」等）；而決策樹是一種基於開發者所決定的分類模型。

分類則是指根據重要的屬性對數據進行分類：這種屬性在機器學習領域被稱為「類別」。

接下來，在步驟3中，使用預先準備的數據集訓練機器學習模型。該任務亦稱為「fitting」。訓練結果的準確性在很大程度上取決於訓練集的內容和數量。

在步驟4，將訓練過的機器學習模型嵌入到系統中。使用者可以利用Pythonlibrary直接運行模型。對於在MCU等元件上運行的機器學習演算法，開發者可以在執行之前將該library轉換為C程式。而針對MEMS MLC等硬體連接型解決方案，開發者可以在執行之前使用專用軟體UNICO-GUI將library轉換為register setting。

最後步驟5，驗證機器學習模型。如果驗證結果與預期的結果不能匹配，則開發者必須審視上述步驟中需要改進的部分，以及如何改進。

類比人類在進行分類的思維方式 – 決策樹

在機器學習興起之前，決策樹模型就已經用於數據分類。例如雜誌上的性格測試 - 在測試中，您需要依照順序回答多個問題（順序是依照每個問題的分支而決定的），之後再根據結果推斷測試者的性格。這是一個典型的決策樹範例。為了創造一個機器學習分類模型，需要在每個階段組合多個檢查點（稱為decision

node)。

在另一個例子中，一位經驗豐富的房地產經紀人評估來訪客戶是否會根據如下圖所示的決策樹購買房產：他將到達一個最終節點（或稱leaf），並根據每個node的條件做出最終決定。

決策樹範例 - 房地產業者

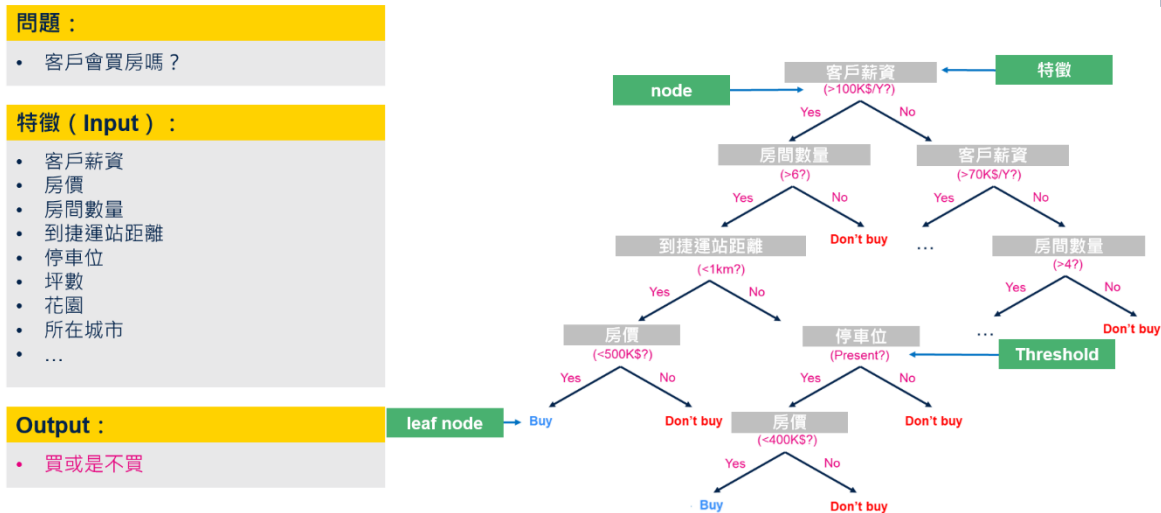


圖3、決策樹範例（以房地產業者為例）

在開發決策樹之前，必須通過分類演算法確定三個要點，即要解決的問題、分類過程中感興趣的參數（特徵值，即input），以及要尋找的最終答案（即output）。在房地產經紀人的例子中，問題是預測客戶是否會購買房產。答案是客戶是否會購買房產。

總之，特徵處理是獲得準確答案的關鍵。在這種情況下，特徵可以包括客戶的收入、房產的價格、房間的數量，以及是否有停車位等等。得到的答案很大程度上取決於特徵值、特徵的順序，以及分類閾值（用於區分數據的thresholds）。為了創造機器學習決策樹，開發者必須先定義需要的特徵和閾值。接著可以透過反覆學習來改善決策樹的結構以獲得更準確的答案，進而最佳化每個node的特徵選擇順序和分類閾值。

決策樹方法看起來像是傳統的「if-then-else」方法，那不同之處在哪呢？

關鍵在於所使用的特徵、決策節點的位置、閾值和決策樹的結構不是由設計人員事先決定的，而是模型根據數據集學習而來。事實上，若依據米蘭、東京或上海的房地產數據集訓練模型（例如，房地產模型）時，可能會得到不同的結果。

[走過AI邊緣裝置的過渡期 – 在內建MLC的ST MEMS感測器上運行嵌入式決策樹。](#)

ST提供了LSM6DSOX、LSM6DSO32X、LSM6DSRX和ISM330DHCX（這些裝置將MEMS感測器與3軸加速度計、3軸陀螺儀和MLC整合在一起），以及IIS2ICLX（整合了超高精度2軸加速度感測器和MLC）。

因應機器學習分類的需求，我們進行了硬體邏輯最佳化且具備出色的感測器規格（圖3），內建MLC的MEMS動作感測器為使用者提供具有低功耗和較高即時精度的個人化嵌入式決策樹（圖3）。MLC可以執行高達32個特徵值的決策樹。意法半導體也提供了開發環境，以便開發者輕鬆開發在MLC中運行的決策樹。

開發者可以在機器學習常用平台（如Weka、RapidMiner、MATLAB和Python（圖4））上進行開發。儘管這些平台相當靈活且多功能，開發者仍需具備足夠的知識和技能才能妥善開發適合該應用的決策樹。

MEMS MLC訓練流程



圖4、MEMS MCL訓練流程

為此，意法半導體針對ST MEMS感測器開發板提供了Unico-GUI軟體，其中包含多種特徵，便於輕鬆開發決策樹並載入至MLC。Unico-GUI使用者能夠輕鬆、直觀地收集數據，並透過簡單易用的GUI來訓練、驗證模型。

意法半導體推出高度通用、配備多種感測器的開發板，可滿足廣泛的硬體規格需求（圖5）。開發板配備支援MLC的MEMS動作感測器，並具有支援無線連接、可由電池供電、可在SD卡上存儲數據等特點，具有高度通用性，可用於廣泛的應用（包括汽車、工業用儀表和消費性電子裝置）。意法半導體更提供AlgoBuilderSuite PC軟體，該軟體（搭配使用Unicleo）提供GUI，可以輕鬆在開發板上操作，並進行數據蒐集。有了AlgoBuilderSuite，開發者可以輕鬆、直觀地對感測器進行程式設計，且無需編寫撰寫任何程式。

數據蒐集 - ST軟體與硬體工具

感測器模組	開發板	專業板
<p>SensorTile.Box STEVAL-MKSBOX1V1</p>  <p>消費者</p>	<p>STM32Nucleo擴充板 X-NUCLEO-IKS01A3</p>  <p>消費者</p>	<p>Profi MEMS工具 STEVAL-MK1109V3</p>  <p>車用 工業用 消費者</p>
<p>STWIN: 無線工業用板 STEVAL-STWINKT1</p>  <p>工業用</p> 	<p>STM32 Nucleo擴充板 X-NUCLEO-IKS02A1</p>  <p>工業用</p> 	<p>軟體解決方案</p> <p>AlgoBuilderSuite</p>    <p>無須撰寫任何一行程式，即可簡單又直觀地透過軟體套件在感測器上進行開發。</p>

圖5、數據收集 – ST軟體與硬體工具