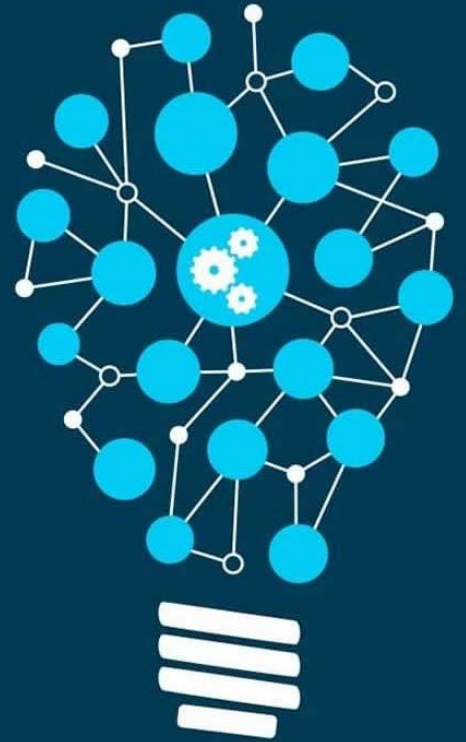


LonGreat ML.NET

No ~~Code~~ Community

ML.NET

MACHINE
LEARNING



張長豐

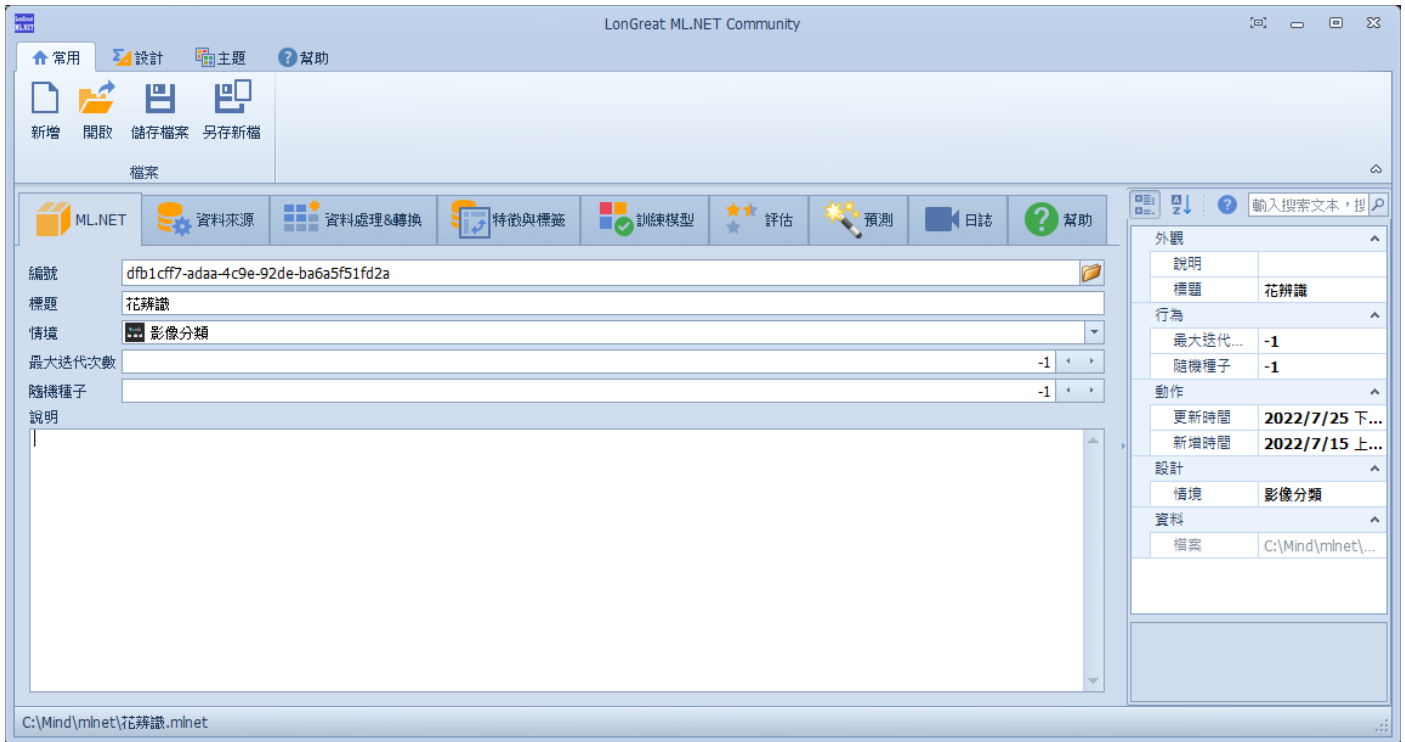
目 錄

內容

簡介.....	2
資料來源.....	4
資料聯合 Federation DataSource	5
檔案目錄轉資料表.....	6
資料處理&轉換.....	7
資料處理程序.....	7
資料轉換函式.....	7
特徵與標籤.....	9
訓練模型.....	10
評估.....	11
預測.....	11
預測應用.....	11

簡介

支援所有資料來源，無須撰寫程式碼



No Code

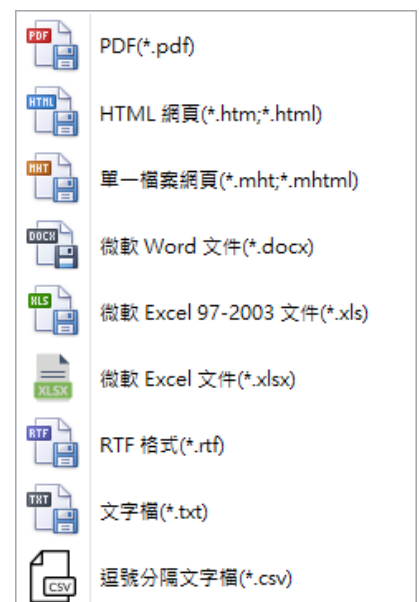
Step by Step

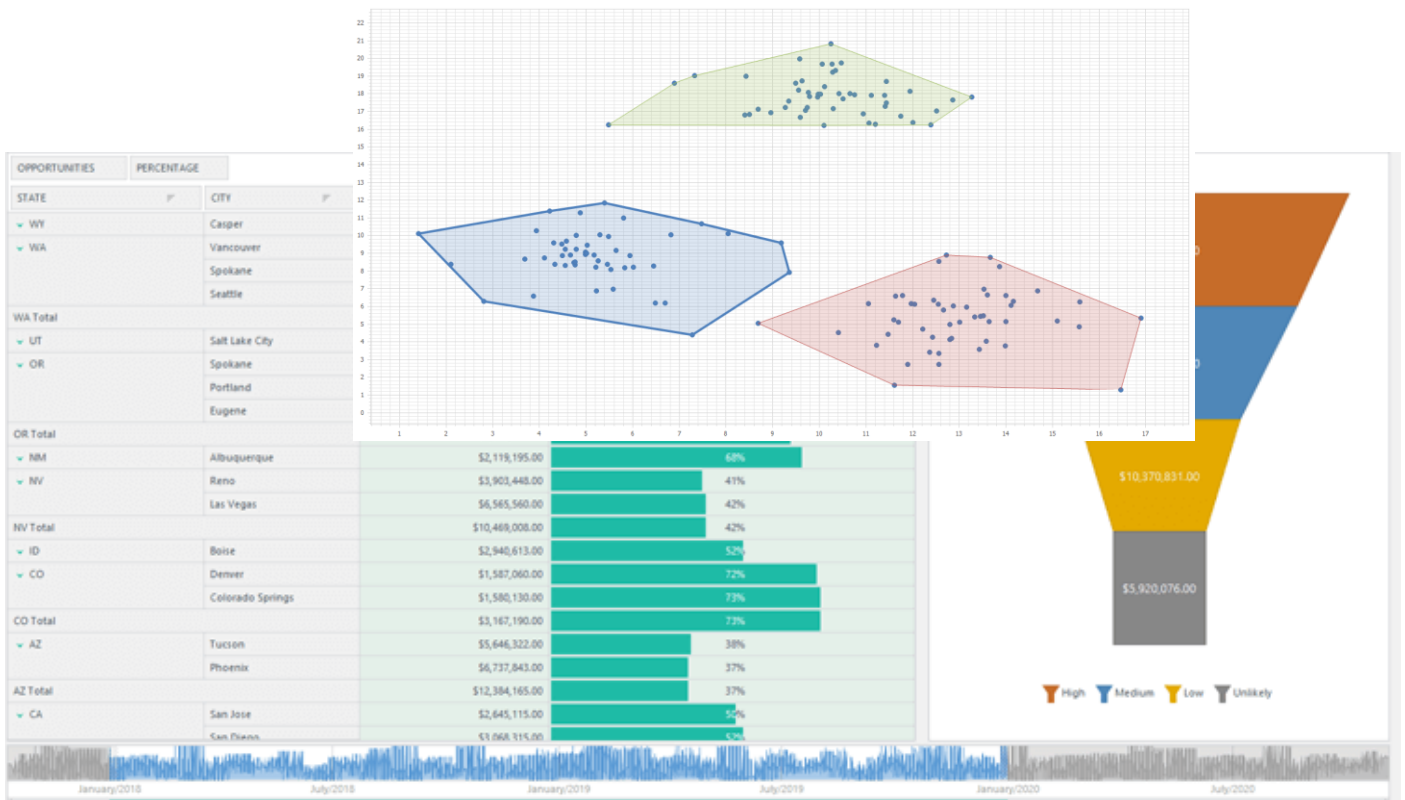
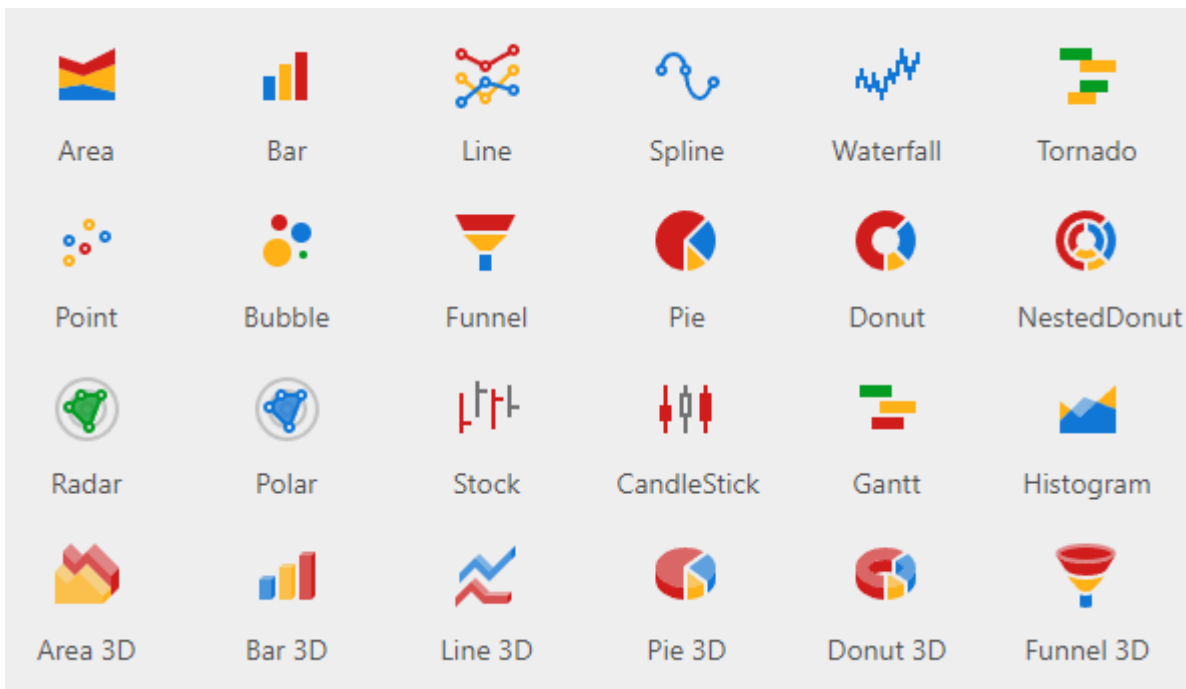
Wizard

Easy & Fast & No Code

Export
diversity

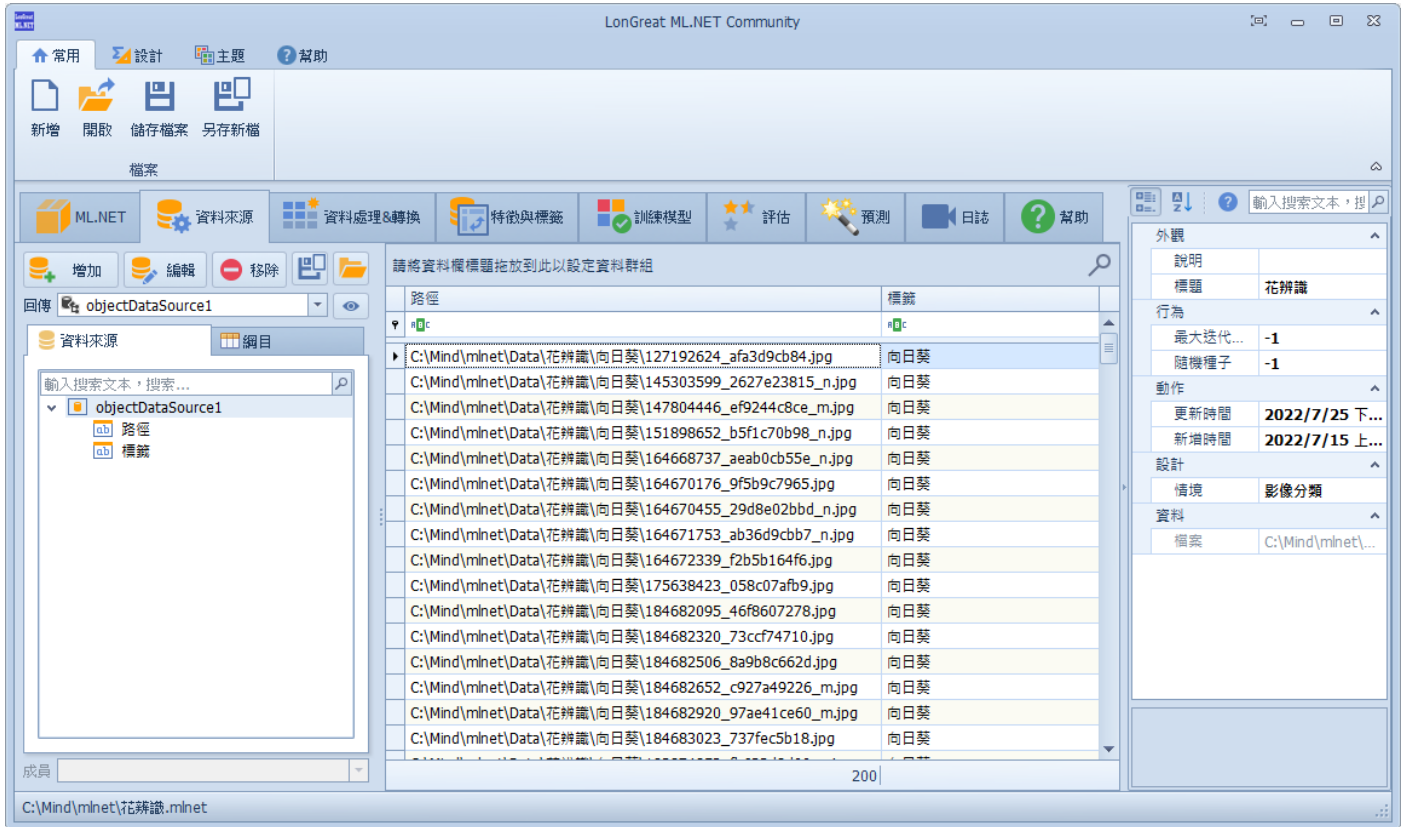
Visualize





資料來源

幾乎涵蓋市面所有資料型態，資料聯合 Federation DataSource 提供不同資料源合併





資料聯合 Federation DataSource

是一個聯合資料源，它集成了不同的數據源並通過聯合查詢提供統一的資料訪問。

建立以下類型的聯合查詢：

Join

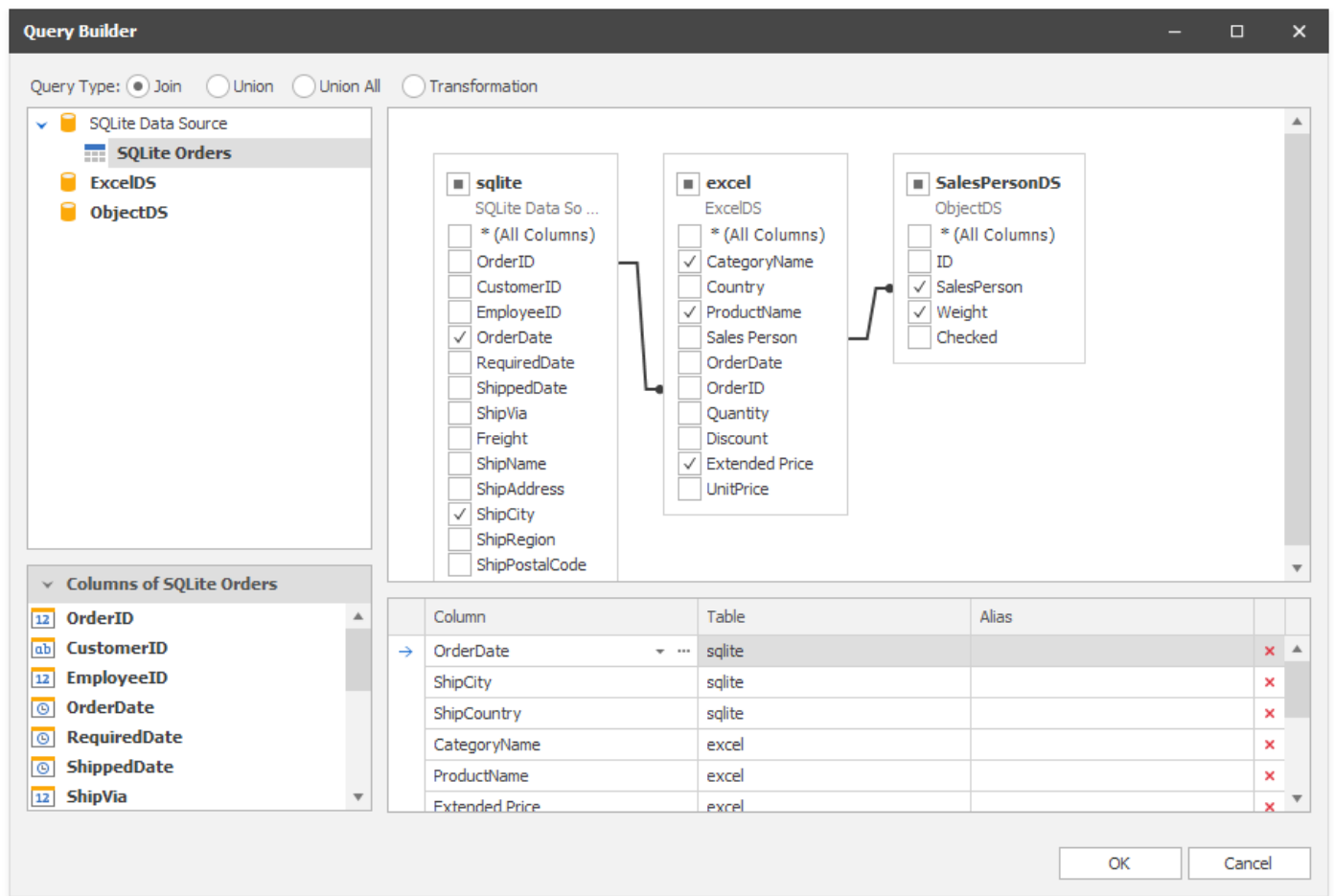
根據共享的列組合來自兩個或多個表的行。連接類型指定在兩個表中具有匹配值的記錄。

Union / UnionAll

Union 查詢將兩個或多個表中的行組合成一個數據集，並刪除合併表中的重複行。UnionAll 查詢的操作類似於 Union，但是當它們包含相同的數據時，它們會復制來自不同表的行。您只能為包含同名列的數據源創建聯合查詢。此類列的數據類型應進行隱式轉換。

Transformation

如果數據源包含複雜的列（對象），您可以轉換其屬性以在平面視圖中將它們顯示為單獨的列。如果其中一個數據列是一個數組，您可以展開它的值並為數組的每個元素顯示一個新的數據行。展開柱時，您可以將其展平並創建展平視圖。



檔案目錄轉資料表



資料處理&轉換

資料處理程序

標題	形別
引導程式	BootstrapSample
快取	Cache
欄位篩選	FilterRowsByColumn
按關鍵欄位分數篩選	FilterRowsByKeyColumnFraction
缺失值欄篩選	FilterRowsByMissingValues
隨機資料列	ShuffleRows
跳過資料列	SkipRows
擷取資料列	TakeRows

資料轉換函式

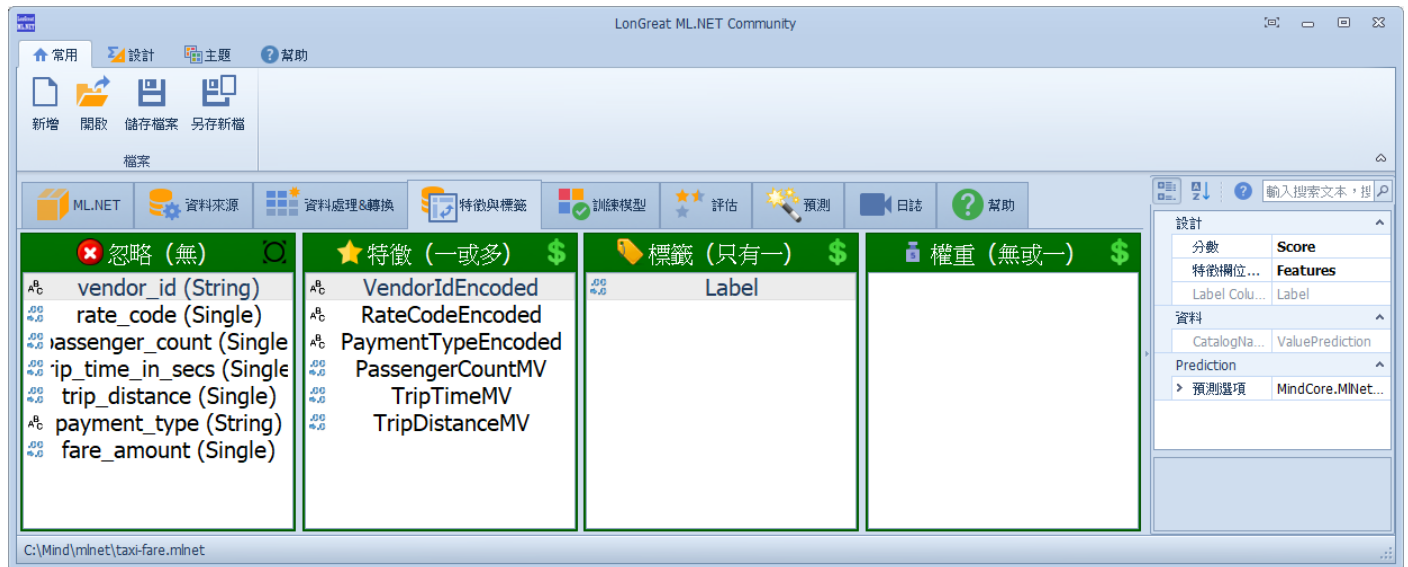
轉換類別	標題	形別
文字轉換	文字內嵌	ApplyWordEmbedding
文字轉換	特徵化文字	FeaturizeText
文字轉換	文字內嵌	LatentDirichletAllocation
文字轉換	標準化文字	NormalizeText
文字轉換	產生雜湊 Ngrams	ProduceHashedNgrams
文字轉換	產生雜湊詞袋	ProduceHashedWordBags
文字轉換	產生 Ngrams	ProduceNgrams
文字轉換	產生詞袋	ProduceWordBags
文字轉換	刪除默認停用詞	RemoveDefaultStopWords
文字轉換	刪除停用詞	RemoveStopWords
文字轉換	將字符標記為鍵	TokenizeIntoCharactersAsKeys
文字轉換	分詞成詞	TokenizeIntoWords
文字轉換	SRCNN 偵測時間異常	DetectAnomalyBySrCnn
文字轉換	通過 Ssa 偵測變化點	DetectChangePointBySsa
文字轉換	偵測變化點	DetectIidChangePoint
文字轉換	偵測 i. i. d 尖峰	DetectIidSpike
文字轉換	Ssa 偵測尖峰	DetectSpikeBySsa

LonGreat ML.NET Community

文字轉換	Ssa 預測	ForecastBySsa
自訂	運算式	DataExpression
核心擴充	近似內核映射	ApproximatedKernelMap
核心擴充	投影到主成分	ProjectToPrincipalComponents
特徵選取	計數選擇特徵	SelectFeaturesBasedOnCount
特徵選取	相互資訊選擇特徵	SelectFeaturesBasedOnMutualInformation
深度學習	簡單 Onnx 模型	SimpleApplyOnnxModel
資料行	連成資料欄	Concatenate
資料行	複製資料欄	CopyColumns
資料行	移除資料行	DropColumns
資料行	載入原始影像	LoadRawImageBytes
資料形別轉換	轉換形別	ConvertType
資料形別轉換	雜湊	Hash
資料形別轉換	鍵轉換二進位	MapKeyToBinaryVector
資料形別轉換	鍵轉換原始值	MapKeyToValue
資料形別轉換	鍵轉換浮點	MapKeyToVector
資料形別轉換	類別轉換鍵	MapValueToKey
標準化	標準 Bin 界線	NormalizeBinning
標準化	標準化全局對比度	NormalizeGlobalContrast
標準化	對數均值方差標準化	NormalizeLogMeanVariance
標準化	標準化調整單位度量	NormalizeLpNorm
標準化	均值方差標準化	NormalizeMeanVariance
標準化	最小值最大值標準化	NormalizeMinMax
標準化	分量範圍調整標準化	NormalizeRobustScaling
標準化	標準化監督分量	NormalizeSupervisedBinning
遺漏值	指示遺漏值	IndicateMissingValues
遺漏值	取代遺漏值	ReplaceMissingValues
類別轉換	最常編碼向量	CategoricalOneHotEncoding
類別轉換	雜湊單一最大編碼向量	CategoricalOneHotHashEncoding

特徵與標籤

拖拉欄位設定



訓練模型

標題	情境	形別
二元平均線性分類模型	二元分類	AveragedPerceptron
二元快速樹系模型	二元分類	FastForestBinary
二元決策樹分類模型	二元分類	FastTreeBinary
欄位感知分解機器模型	二元分類	FieldAwareFactorizationMachine
二元一般化加法模型	二元分類	GamBinary
二元邏輯回歸模型	二元分類	LbfgsLogisticRegressionBinary
二元促進決策樹模型	二元分類	LightGbmBinary
二元線性 SVM 分類模型	二元分類	LinearSvm
二元分類模型	二元分類	Prior
二元邏輯回歸分類模型	二元分類	SdcaLogisticRegressionBinary
二元布林回歸分類模型	二元分類	SdcaNonCalibratedBinary
二元符號隨機梯度下降	二元分類	SymbolicSgdLogisticRegressionBinary
單變數時間序列模型	預測	ForecastBySsa
DNN 影像分類模型	影像分類	ImageClassification
矩陣分解模型	建議	MatrixFactorization
隨機 SVD 模型	異常偵測	RandomizedPca
多元 L BFGS 最大熵模型	多類分類	LbfgsMaximumEntropy
多元 LightGBM 促進式決策樹模型	多類分類	LightGbm
多元貝氏機率分類模型	多類分類	NaiveBayes
最大熵多元分類模型	多類分類	SdcaMaximumEntropy
線性多元分類模型	多類分類	SdcaNonCalibrated
快速決策樹回歸模型	值預測	FastForestRegression
決策樹回歸模型	值預測	FastTreeRegression
複合波氏決策樹回歸模型	值預測	FastTreeTweedie
一般化加法回歸模型	值預測	GamRegression
波氏回歸模型	值預測	LbfgsPoissonRegression
促進決策樹回歸模型	值預測	LightGbmRegression
最小平方回歸模型	值預測	OlsRegression
梯度下降回歸模型	值預測	OnlineGradientDescent
隨機雙重座標回歸模型	值預測	SdcaRegression
K-means 分群	群集	KMeans

評估

評估紀錄與相關分數紀錄

LonGreat ML.NET Community

常用 設計 主題 幫助

自動新增模型 預覽列印 匯出 模型資料夾

模型 檢視 模型

ML.NET 資料來源 資料處理 & 轉換 特徵與標籤 訓練模型 評估 預測 日誌 幫助

評估 網絡顯示

請將資料欄標題拖放到此以設定資料群組

編號	資料來源	筆數	編號	開始時間	結束時間	時間	指標
2022-07-21-0...	DataSo...	0	2022-07-21-09-08-22 ...	2022年7月21日 上...	2022年7月21日 ...	00:00:02...	

資料

平均絕對誤差 (MAE)	0.8756125531148911
判定係數	0.6380906389490668
均方誤差 (MSE)	42.38244664073625
根均方誤差 (RMSE)	6.5101802310486185
損失函數	42.38244700233495

判定係數

R-Squared 接近 1 表示模型擬合度更好。

C:\Mind\mlnet\taxi-fare.minet

預測

批次預測與輸入預測

預測應用

訓練模型產生 zip 檔，程式載入使用