

Community V1.0.0

目 录

内容

[简介 2](#_Toc110407638)

[支持所有数据源，无须撰写程序代码 2](#_Toc110407639)

[多种格式汇出与打印 3](#_Toc110407640)

[强大图表 3](#_Toc110407641)

[数据透视表与连动图表 4](#_Toc110407642)

[资料网格检视 5](#_Toc110407643)

[数据源 6](#_Toc110407644)

[数据联合Federation DataSource 7](#_Toc110407645)

[档案目录转数据表 8](#_Toc110407646)

[数据处理&转换 9](#_Toc110407647)

[数据处理程序 9](#_Toc110407648)

[数据转换函式 9](#_Toc110407649)

[特征与卷标 11](#_Toc110407650)

[训练模型 12](#_Toc110407651)

[简单与专业属性 13](#_Toc110407652)

[评估 14](#_Toc110407653)

[预测 14](#_Toc110407654)

[预测应用 14](#_Toc110407655)

[使用相关技术 14](#_Toc110407656)

[相关连结 15](#_Toc110407657)

# 简介

## 支持所有数据源，无须撰写程序代码

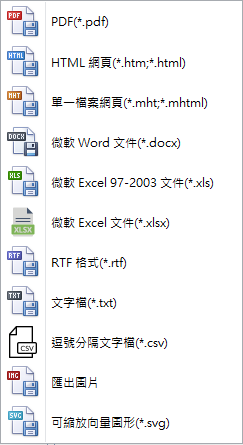
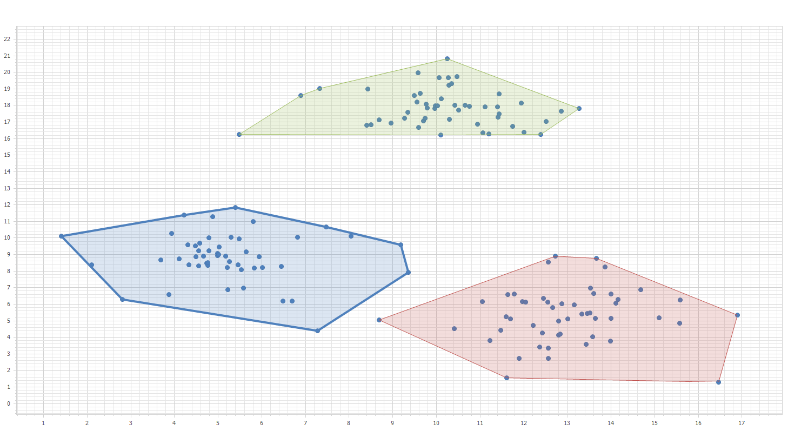
一張含有 文字 的圖片

自動產生的描述

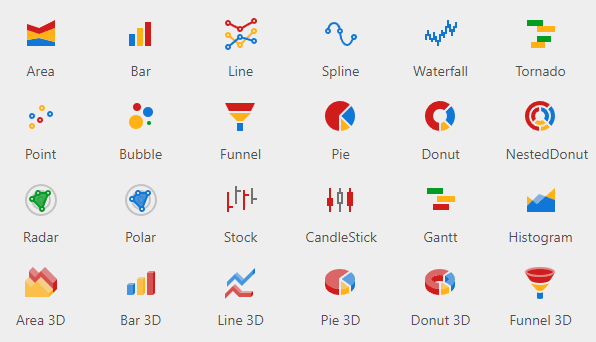


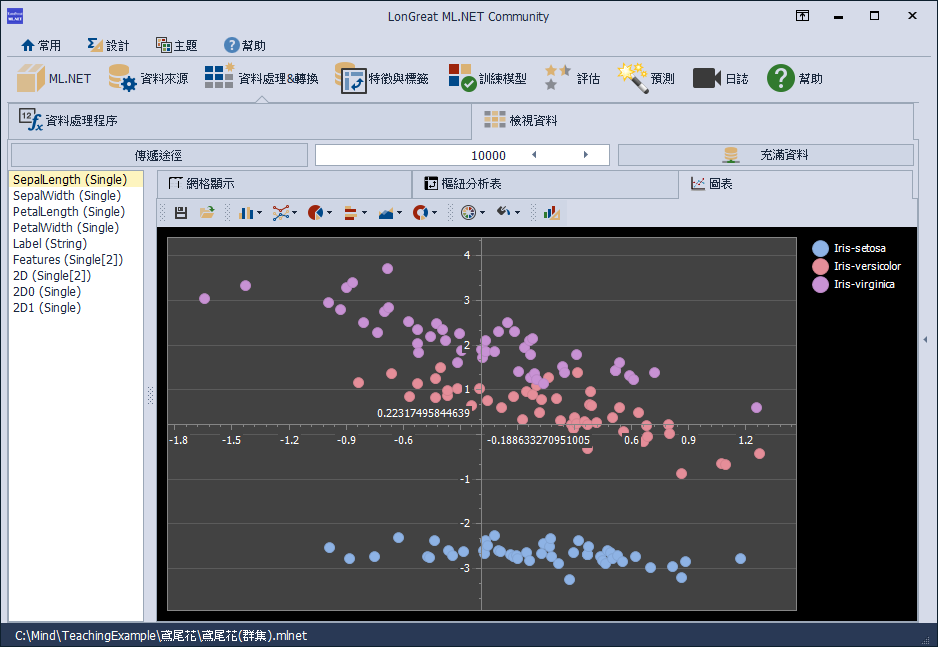
## **多种格式汇出与打印**

所见既所得打印与汇出



## 强大图表

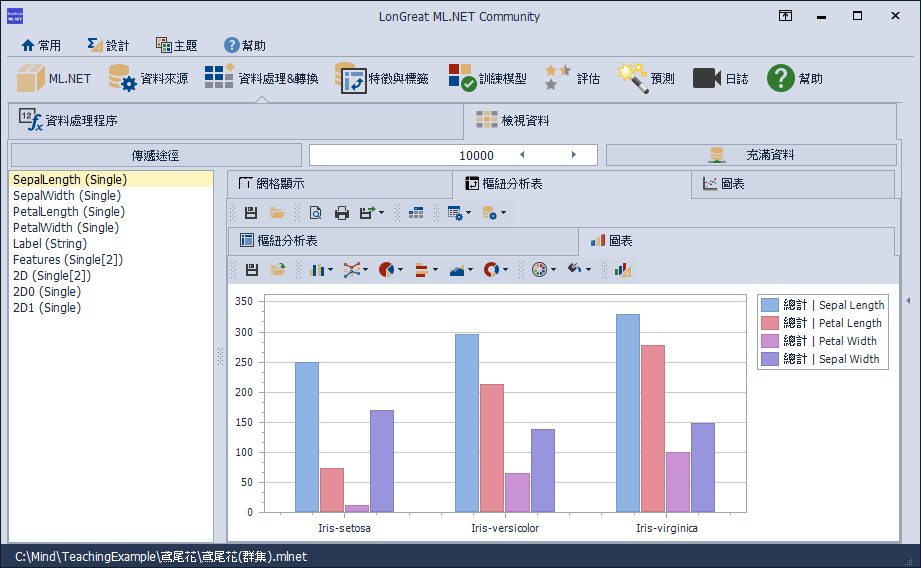




数据透视表与连动图表

一張含有 文字 的圖片

自動產生的描述



## 资料网格检视

快速、群组、筛选、、、

一張含有 桌 的圖片

自動產生的描述

# 数据源

几乎涵盖市面所有数据型态，数据联合Federation DataSource提供不同数据源合并

一張含有 文字 的圖片

自動產生的描述

一張含有 文字 的圖片

自動產生的描述

## 数据联合Federation DataSource

是一个联合数据源，它集成了不同的数据源并通过联合查询提供统一的数据访问。

建立以下类型的联合查询：

**Join**

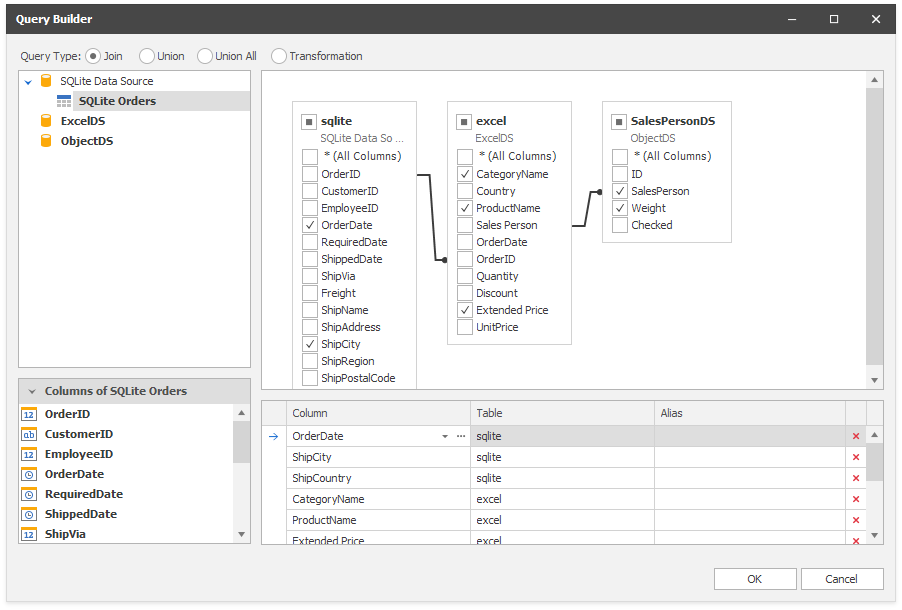
根据共享的列组合来自两个或多个表的行。连接类型指定在两个表中具有匹配值的记录。

Union / UnionAll

Union查询将两个或多个表中的行组合成一个数据集，并删除合并表中的重复行。UnionAll查询的操作类似于Union ，但是当它们包含相同的数据时，它们会复制来自不同表的行。您只能为包含同名列的数据源创建联合查询。此类列的数据类型应进行隐式转换。

**Transformation**

如果数据源包含复杂的列（对象），您可以转换其属性以在平面视图中将它们显示为单独的列。如果其中一个数据列是一个数组，您可以展开它的值并为数组的每个元素显示一个新的数据行。展开柱时，您可以将其展平并创建展平视图。



## 档案目录转数据表

一張含有 桌 的圖片

自動產生的描述

# 数据处理&转换

## 数据处理程序

|  |  |
| --- | --- |
| 标题 | 形别 |
| 引导程序 | BootstrapSample |
| 快取 | Cache |
| 字段筛选 | FilterRowsByColumn |
| 按关键词段分数筛选 | FilterRowsByKeyColumnFraction |
| 缺失值栏筛选 | FilterRowsByMissingValues |
| 随机资料列 | ShuffleRows |
| 跳过资料列 | SkipRows |
| 撷取资料列 | TakeRows |

## 数据转换函式

|  |  |  |
| --- | --- | --- |
| 转换类别 | 标题 | 形别 |
| 文字转换 | 文字内嵌 | ApplyWordEmbedding |
| 文字转换 | 特征化文字 | FeaturizeText |
| 文字转换 | 文字内嵌 | LatentDirichletAllocation |
| 文字转换 | 标准化文字 | NormalizeText |
| 文字转换 | 产生哈希Ngrams | ProduceHashedNgrams |
| 文字转换 | 产生哈希词袋 | ProduceHashedWordBags |
| 文字转换 | 产生Ngrams | ProduceNgrams |
| 文字转换 | 产生词袋 | ProduceWordBags |
| 文字转换 | 删除默认停用词 | RemoveDefaultStopWords |
| 文字转换 | 删除停用词 | RemoveStopWords |
| 文字转换 | 将字符标记为键 | TokenizeIntoCharactersAsKeys |
| 文字转换 | 分词成词 | TokenizeIntoWords |
| 文字转换 | SRCNN侦测时间异常 | DetectAnomalyBySrCnn |
| 文字转换 | 通过 Ssa 侦测变化点 | DetectChangePointBySsa |
| 文字转换 | 侦测变化点 | DetectIidChangePoint |
| 文字转换 | 侦测i.i.d尖峰 | DetectIidSpike |
| 文字转换 | Ssa侦测尖峰 | DetectSpikeBySsa |
| 文字转换 | Ssa预测 | ForecastBySsa |
| 自定义 | 表达式 | DataExpression |
| 核心扩充 | 近似核心对应 | ApproximatedKernelMap |
| 核心扩充 | 投影到主成分 | ProjectToPrincipalComponents |
| 特征选取 | 计数选择特征 | SelectFeaturesBasedOnCount |
| 特征选取 | 相互信息选择特征 | SelectFeaturesBasedOnMutualInformation |
| 深度学习 | 简单Onnx模型 | SimpleApplyOnnxModel |
| 资料行 | 连成数据域 | Concatenate |
| 资料行 | 复制数据域 | CopyColumns |
| 资料行 | 移除资料行 | DropColumns |
| 资料行 | 加载原始影像 | LoadRawImageBytes |
| 数据形别转换 | 转换形别 | ConvertType |
| 数据形别转换 | 哈希 | Hash |
| 数据形别转换 | 键转换二进制 | MapKeyToBinaryVector |
| 数据形别转换 | 键转换原始值 | MapKeyToValue |
| 数据形别转换 | 键转换浮点 | MapKeyToVector |
| 数据形别转换 | 类别转换键 | MapValueToKey |
| 标准化 | 标准Bin界线 | NormalizeBinning |
| 标准化 | 标准化全局对比度 | NormalizeGlobalContrast |
| 标准化 | 对数均值方差标准化 | NormalizeLogMeanVariance |
| 标准化 | 标准化调整单位度量 | NormalizeLpNorm |
| 标准化 | 均值方差标准化 | NormalizeMeanVariance |
| 标准化 | 最小值最大值标准化 | NormalizeMinMax |
| 标准化 | 分量范围调整标准化 | NormalizeRobustScaling |
| 标准化 | 标准化监督分量 | NormalizeSupervisedBinning |
| 遗漏值 | 指示遗漏值 | IndicateMissingValues |
| 遗漏值 | 取代遗漏值 | ReplaceMissingValues |
| 类别转换 | 最常编码向量 | CategoricalOneHotEncoding |
| 类别转换 | 哈希单一最大编码向量 | CategoricalOneHotHashEncoding |

# 特征与卷标

拖拉字段设定

一張含有 桌 的圖片

自動產生的描述

# 训练模型

|  |  |  |
| --- | --- | --- |
| 标题 | 情境 | 形别 |
| 二元平均线性分类模型 | 二元分类 | AveragedPerceptron |
| 二元快速树系模型 | 二元分类 | FastForestBinary |
| 二元判定树分类模型 | 二元分类 | FastTreeBinary |
| 字段感知分解机器模型 | 二元分类 | FieldAwareFactorizationMachine |
| 二元一般化加法模型 | 二元分类 | GamBinary |
| 二元逻辑回归模型 | 二元分类 | LbfgsLogisticRegressionBinary |
| 二元促进判定树模型 | 二元分类 | LightGbmBinary |
| 二元线性SVM分类模型 | 二元分类 | LinearSvm |
| 二元分类模型 | 二元分类 | Prior |
| 二元逻辑回归分类模型 | 二元分类 | SdcaLogisticRegressionBinary |
| 二元布尔回归分类模型 | 二元分类 | SdcaNonCalibratedBinary |
| 二元符号随机梯度下降 | 二元分类 | SymbolicSgdLogisticRegressionBinary |
| 单变量时间序列模型 | 预测 | ForecastBySsa |
| DNN影像分类模型 | 影像分类 | ImageClassification |
| 矩阵分解模型 | 建议 | MatrixFactorization |
| 随机SVD模型 | 异常侦测 | RandomizedPca |
| 多元 L BFGS 最大熵模型 | 多类分类 | LbfgsMaximumEntropy |
| 多元 LightGBM 促进式判定树模型 | 多类分类 | LightGbm |
| 多元贝氏机率分类模型 | 多类分类 | NaiveBayes |
| 最大熵多元分类模型 | 多类分类 | SdcaMaximumEntropy |
| 线性多元分类模型 | 多类分类 | SdcaNonCalibrated |
| 快速判定树回归模型 | 值预测 | FastForestRegression |
| 判定树回归模型 | 值预测 | FastTreeRegression |
| 复合波氏判定树回归模型 | 值预测 | FastTreeTweedie |
| 一般化加法回归模型 | 值预测 | GamRegression |
| 波氏回归模型 | 值预测 | LbfgsPoissonRegression |
| 促进判定树回归模型 | 值预测 | LightGbmRegression |
| 最小平方回归模型 | 值预测 | OlsRegression |
| 梯度下降回归模型 | 值预测 | OnlineGradientDescent |
| 随机双重坐标回归模型 | 值预测 | SdcaRegression |
| K-means 分群 | 群集 | KMeans |

## 简单与专业属性

训练器提供简单与专业两种选择

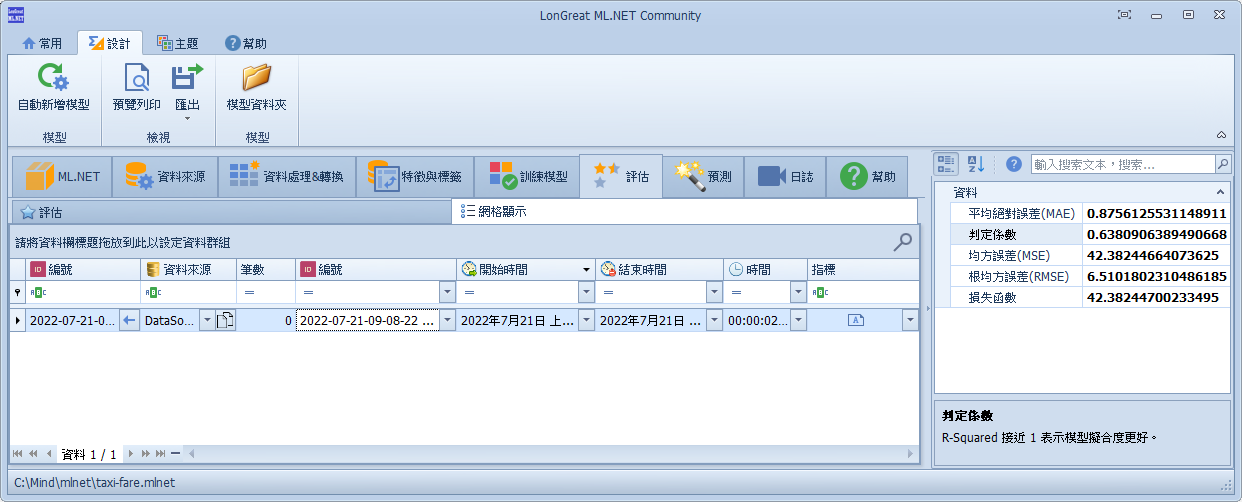
一張含有 桌 的圖片

自動產生的描述 一張含有 桌 的圖片

自動產生的描述

# 评估

评估纪录与相关分数纪录



# 预测

批次预测与输入预测

# 预测应用

训练模型产生zip文件，程序加载发布应用

# 使用相关技术

[NET6](https://docs.microsoft.com/dotnet/core/whats-new/dotnet-6)

[ML.NET 1.7.1](https://docs.microsoft.com/dotnet/machine-learning/tutorials/)

[Devexpress](https://www.devexpress.com/)

# 相关连结

|  |  |
| --- | --- |
| GitHub 入門(一) — 如何建立Git專案. 引言| by 趙子榮| Medium | [github.longreat.net](http://github.longreat.net/) |
| 关于我们- Gitee.com | [gitee.longreat.net](http://gitee.longreat.net/) |