PMML是一种数据挖掘模型的标记定义语言，业界统一较权威。

通过将挖掘模型转换为PMML，可以使得不同挖掘软件之间模型的共享。

以指数平滑算法生成的模型为例，介绍PMML标记语言的语法结构。

（此PMML来源于<http://www.dmg.org/>）

**# PMML版本。**

<PMML xmlns="http://www.dmg.org/PMML-4\_1" version="4.1">

# PMML 头，定义应用名称，生成时间戳。

<Header copyright="DMG.org">

<Application name="test application"/>

<Timestamp>2008-06-23 10:30:00</Timestamp>

</Header>

**# 挖掘模型数据字典，对模型输入数据的定义，比如次例中，输入数据为TS，Value两个字段，TS类似与日期索引，Value为其对应的指标值。**

<DataDictionary numberOfFields="2">

<DataField dataType="integer" optype="continuous" name="TS" displayName="TS"/>

<DataField dataType="double" optype="continuous" name="VALUE" displayName="TS-VALUE"/>

</DataDictionary>

**# 开始定义挖掘模型，定义模型，函数，算法。**

<TimeSeriesModel modelName="AA2Model" functionName="timeSeries" algorithmName="exponential smoothing" bestFit="ExponentialSmoothing">

**#挖掘模式，将输入数据与算法的相关角色对应起来，比如Value字段为预测字段。**

<MiningSchema>

<MiningField name="TS" usageType="order"/>

<MiningField name="VALUE" usageType="predicted"/>

</MiningSchema>

**#对输入时间序列数据的定义，24个时间序列值，用于训练模型。**

<TimeSeries usage="logical" startTime="1" endTime="24" interpolationMethod="none">

<TimeValue index="1" value="112"/>

<TimeValue index="2" value="118"/>

<TimeValue index="3" value="132"/>

<TimeValue index="4" value="129"/>

<TimeValue index="5" value="121"/>

<TimeValue index="6" value="135"/>

<TimeValue index="7" value="148"/>

<TimeValue index="8" value="148"/>

<TimeValue index="9" value="136"/>

<TimeValue index="10" value="119"/>

<TimeValue index="11" value="104"/>

<TimeValue index="12" value="118"/>

<TimeValue index="13" value="115"/>

<TimeValue index="14" value="126"/>

<TimeValue index="15" value="141"/>

<TimeValue index="16" value="135"/>

<TimeValue index="17" value="125"/>

<TimeValue index="18" value="149"/>

<TimeValue index="19" value="170"/>

<TimeValue index="20" value="170"/>

<TimeValue index="21" value="158"/>

<TimeValue index="22" value="133"/>

<TimeValue index="23" value="114"/>

<TimeValue index="24" value="140"/>

</TimeSeries>

**#预测结果集。**

<TimeSeries usage="prediction" interpolationMethod="none">

<TimeValue index="25" value="145" standardError="7.3"/>

<TimeValue index="26" value="150" standardError="8.3"/>

<TimeValue index="27" value="178" standardError="9.3"/>

<TimeValue index="28" value="163" standardError="10.3"/>

<TimeValue index="29" value="172" standardError="11.3"/>

<TimeValue index="30" value="178" standardError="12.3"/>

<TimeValue index="31" value="199" standardError="13.3"/>

<TimeValue index="32" value="199" standardError="14.3"/>

<TimeValue index="33" value="184" standardError="15.3"/>

<TimeValue index="34" value="162" standardError="16.3"/>

<TimeValue index="35" value="146" standardError="17.3"/>

<TimeValue index="36" value="166" standardError="18.3"/>

</TimeSeries>

**#指数平滑模型参数，模型标准误差7.3。**

<ExponentialSmoothing RMSE="7.3">

**# 平滑常数为0.233984，平滑值为139.8。**

<Level alpha="0.233984" smoothedValue="139.8"/>

**#趋势分析，平滑值4.139，趋势模型为damped\_additive，及其各种参数值。**

<Trend\_ExpoSmooth smoothedValue="4.139" gamma="3.910E-4" phi="1.006" trend="damped\_additive"/>

**#季节分析，以12个月为一个周期（即一年），使用季节乘法模型。**（type="multiplicative"），及模型的各参数。

<Seasonality\_ExpoSmooth type="multiplicative" period="12" unit="month" delta="0.8254" phase="12">

**#季节乘法模型中每个月份的季节比率。**

<Array n="12" type="real">

.900 .840 .924 .976 .994 1.120 0.981 1.025 1.038 1.038 0.908 1.259

</Array>

</Seasonality\_ExpoSmooth>

</ExponentialSmoothing>

</TimeSeriesModel>

</PMML>