Программа кластеризации данных существует и способна функционировать отдельно.  
“ImportsScrolls”- класс контейнер классов взаимодействия с другими программами.  
(По одному классу на программу)  
ImportsScrolls.FuzzyForecastImportScroll – класс взаимодействия программы “ClusterIII” с программой “FuzzyForecast”

“ImportsScrolls.FuzzyForecastImportScroll” содержит в себе методы взаимодействия программы “FuzzyForecast” с программой “ClusterIII”.

Общий принцип работы методов.

В программе экспортёре (“FuzzyForecast”) вызывается один из методов взаимодействия и передаются в него данные. В методе взаимодействия прописывается сценарий.

Например в методе “ImportsScrolls.FuzzyForecastImportScroll. Scroll2(List<List<double>> DataSet, List<string> DataNameSet, List<string> DataTypeSet, List<string> DataKSet)”  
описывается

---------------------------------------------------------------------------------------------------------------------------------------

***Предназначение сценария:***  
-Обнаружение аномалий временного ряда при помощи центроидного метода кластеризации значений ряда.  
-Сглаживание обнаруженных аномалий временного ряда.

***Особенности сценария:***  
-Обрабатываются все загруженные файлы-ряда в “FuzzyForecast”.   
(“Загрузил, дал задание и забыл”)

-За процессом обработки можно следить при помощи окон индикаторных полос прогресса.  
-Пред обработкой заранее спрашивается папка сохранения файлов.  
-Значительное время обработки, которое может быть сокращено неоднократным рефакторингом кода программы. (Есть к чему стремиться)

***Этапы сценария:***

1)Создание рабочей предметной области для кластеризации

Сохранение всех обрабатываемых данных в виде типа данных “Cluster”

2)Получение результатов кластеризации в виде таблицы в виде типа данных “DataTable”

Используютется программа “ClusterIII”. Входные данные типа “Cluster”.

Кластеризация центройдным методом на 2 кластера. В каждом кластере остаётся минимум один объект.

3)Получение конечных кластеров в виде соответствующих списков объектов в виде типа данных “List<string>”

4)Поиск самого малочисленного кластера в виде соответствующего ему списка объектов в виде типа данных “List<string>”

Считаем что самый малочисленный (из двух) список – список аномалий.

Считаем что априори в ряде есть хотя бы одна аномалия. Иначе зачем запускать данный сценарий.

5)Сглаживание аномалий в соответствии с полученным списком.

Находится аномалия и сглаживается по формуле

Где – значение ряда, i=0..n, m - радиус окрестности аномалии. В нашем случае m=3;

DataSet[i][j] = (Int64)

(DataSet[i][j - 3] + DataSet[i][j - 2]+ DataSet[i][j - 1]+

DataSet[i][j - 1]+ DataSet[i][j - 2] +DataSet[i][j + 3])

/(2\*3);

6)Сохранение результатов в файл по заранее указанному пути

Берётся имя файла ряда , загруженного в “FuzzyForecast”, добавляется “\_RemoveAnomalies”.

Перспективы модернизации:

1. Рефакторингом
2. Обработка не самого ряда, а отклонений его значений от тренда или заранее сглаженного ряда (возможно методом скользящей средней)
3. Итеративное сглаживание ряда, сглаживание условных аномалий, обнаруженных на каждой итерации.