

## ОСОБЕННОСТИ ПРОЦЕССА ВЫЯВЛЕНИЯ АНОМАЛЬНЫХ ЗНАЧЕНИЙ В СИСТЕМЕ WOLFRAM MATHEMATICA

*Каменко Д.А.,*

*магистрант ПСФ, УО «БНТУ», г. Минск, Республика Беларусь*

*Научный руководитель – Гундина М.А., канд. физ.-мат. наук, доцент*

Ключевые слова. Аномальные значения, выборка, компьютерная система Wolfram Mathematica, статистическая совокупность.

Keywords. Anomalous values, sampling, Wolfram Mathematica computer system, statistical universe.

Задача автоматизации процесса выявления аномальных значений выборки не теряет свою актуальность в инженерной практике. Известно, что аномальные значения способны существенно исказить функционирование математических моделей анализа данных [1], что может привести к снижению надежности и некорректной работе всей системы. Единицы статистической совокупности, у которых значения анализируемого признака существенно отклоняются от основного массива, называются аномальными значениями [2].

**Материал и методы.** В качестве компьютерной системы для организации процесса автоматизации определения аномальных значений выбрана система Wolfram Mathematica. Она позволяет применять методы машинного обучения решения данной задачи.

**Результаты и их обсуждение.** Функция системы FindAnomalies позволяет найти члены выборки, которые считаются аномальными по отношению к выборке. Задавая порог принятия можно в автоматическом режиме определять аномальные значения выборки. Рассмотрим пример кода, определяющего аномальные значения:

```
ndata=RandomReal[1,{100,3}];  
adata=RandomReal[{-2,-0.5},{5,3}];  
data=Join[ndata,adata];
```

```
ListPointPlot3D[{ndata,FindAnomalies[data,AcceptanceThreshold->0.1]}].
```

Результат работы кода представлен на рисунке 1.

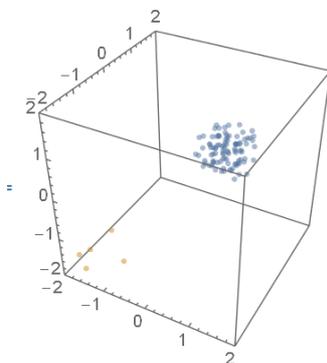


Рисунок 1 – Автоматическое определение аномальных значений

Данная функция может быть использована для многих типов данных, включая числовые, строковые и графические данные.

Рассмотрим задачу определение аномальных температур, наблюдаемых в Минске за некоторый промежуток времени.

```
forecast =WeatherData["Minsk","MeanTemperature",{2017,1,1},{2017,12,31},"Day"]  
tsp=forecast["Path"];
```

Сформируем массив, первая координата точек которого будет определять день по порядку, вторая координата – соответствующую наблюдаемую температуру:

```
For[tsp1=tsp;i=1,i<=Length[tsp],i++,tsp1[[i]]={i,forecast["Path"][[i,2]]}]
```

График распределения температуры представлен на рисунке 2.

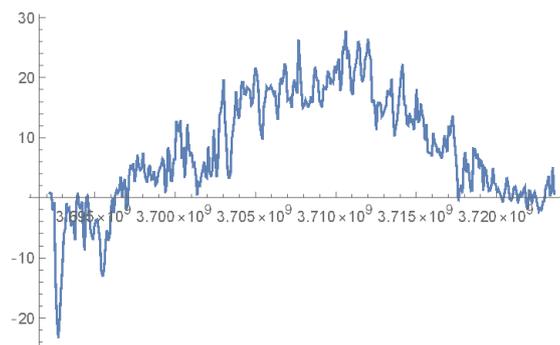


Рисунок 2 – График распределения температуры

Определим аномальные значения массива:

```
anomalies=FindAnomalies[tsp1]
```

Результат выполнения функции выглядит следующим образом:

```
Out[514]= {{6, 0.92278 C°},{7, 0.911253 C°},{8, 0.923181 C°}}
```

В системе также есть и другие функции, позволяющие анализировать аномальные значения. Функция DeleteAnomalies выдает новый набор данных, в котором уже удалены аномальные значения. Рассмотрим некоторую выборку, содержащую аномальные значения:

```
data={5,10,6,100,7,8,7,6,10,9,4,5,8,5};
```

Построим график исходных данных и данных, в которых уже отсутствуют аномальные значения (рисунок 3):

```
ListLinePlot[data, PlotRange->All]
```

```
ListLinePlot[DeleteAnomalies[data]]
```

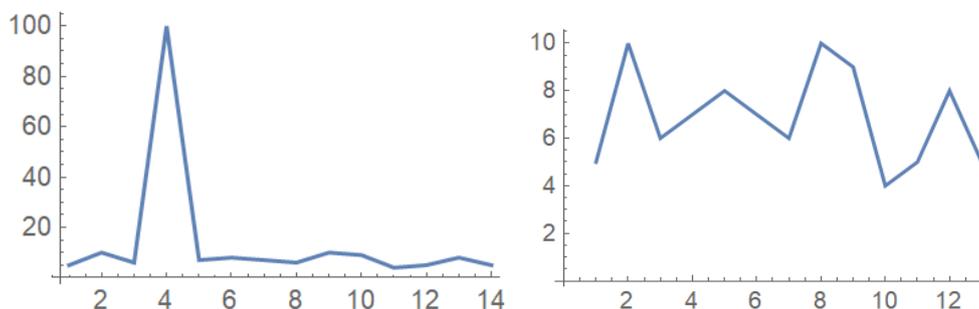


Рисунок 3 – График исходных значений. График значений без аномалий

Функции AnomalyDetection и AnomalyDetectorFunction позволяют проверить, является ли новое предъявляемое значение аномальным.

**Заключение.** Процесс выявления аномалий является очень важным вопросом в задачах предсказания поломок оборудования, выявления аномального спроса на потребляемую продукцию и др.

1. Dau, H. Anomaly Detection Using Replicator Neural Networks Trained on Examples of One Class / H. Dau, V. Ciesielski, A. Song // Simulated Evolution and Learning. Lecture Notes in Computer Science. – 2014. – V. 8886. – P. 311–322.

2. Network anomaly detection using a fuzzy rule-based classifier / S. Ghosh [et al.] // Computer, Communication and Electrical Technology. – 2017. – P. 61–65.