



Скользящее окно

Обработка данных методом скользящего окна применяется при предварительной обработке данных в задачах прогнозирования, когда на вход анализатора (например, нейронной сети) требуется подавать значения нескольких смежных отсчетов исходного набора данных. Термин скользящее окно отражает сущность обработки — выделяется некоторый непрерывный отрезок данных, называемый окном, а окно, в свою очередь, перемещается, "скользит" по всему исходному набору данных.

В результате будет получен набор данных, где в одном поле будет содержаться значение, соответствующее текущему отсчету (оно будет иметь то же имя, что и в исходном наборе), а слева и справа от него будут расположены поля со значениями, смещенными от текущего отсчета в прошлое и в будущее соответственно.

Следовательно, обработка методом скользящего окна имеет два параметра:

- **Глубина истории** — количество отсчетов в "прошлое";
- **Горизонт прогноза** — количество отсчетов в "будущее".

Необходимо отметить, что для граничных положений окна (конец и начало исходной выборки) будут формироваться неполные записи: в начале исходной выборки будут формироваться пустые значения для "прошлых" отсчетов, а в конце — для "будущих". В зависимости от конкретной ситуации пользователь может включать такие неполные записи в результирующую выборку или исключать их.

▼ Пример:

Исходная таблица:

Дата	Продажи, шт.
01.09.2011	45
01.10.2011	82
01.11.2011	120
01.12.2011	192
01.01.2012	229
01.02.2012	161

Для поля `продажи, шт.` настроим параметр *Глубина истории* равным двум, а параметр *Горизонт прогноза* — равным единице. В зависимости от параметра *Оставлять неполные записи* получим разные результирующие таблицы.

Результирующая таблица при значении *Оставлять неполные записи*:

Дата	Продажи, шт. [-2]	Продажи, шт. [-1]	Продажи, шт.	Продажи, шт. [+1]
				45
01.09.2011			45	82
01.10.2011		45	82	120
01.11.2011	45	82	120	192
01.12.2011	82	120	192	229
01.01.2012	120	192	229	161
01.02.2012	192	229	161	
	229	161		
	161			

Порты

Вход

-  **Входной источник данных** — порт для подключения входного набора данных.

Выход

-  **Выходной набор данных** — на порт выводится таблица с набором данных дополненным смещенными полями.

Мастер настройки

Окно мастера настройки содержит список полей входной таблицы, для каждого поля имеются настраиваемые параметры:

- **Глубина истории** — количество значений из предыдущих записей, для которых создаются новые поля в выходном наборе данных;
- **Горизонт прогноза** — количество значений из последующих записей, для которых создаются новые поля в выходном наборе данных.

Параметр *Способ обработки неполных записей* предоставляет следующие методы:

- **Оставлять неполные записи** — сохраняет все добавленные узлом записи;
- **Удалять добавленные неполные записи** — удаляет записи добавленные узлом, не трогая записи из изначального набора;
- **Удалять все неполные записи** — удаляет записи добавленные узлом и записи с пустыми значениями в добавленных полях.

Варианты результирующей таблицы из примера с разными *Способами обработки неполных записей*.

Результирующая таблица при значении *Оставлять неполные записи*:

Дата	Продажи, шт.[-2]	Продажи, шт.[-1]	Продажи, шт.	Продажи, шт.[+1]
				45
01.09.2011			45	82
01.10.2011		45	82	120
01.11.2011	45	82	120	192
01.12.2011	82	120	192	229
01.01.2012	120	192	229	161
01.02.2012	192	229	161	
	229	161		
	161			

Результирующая таблица при значении *Удалять добавленные неполные записи*:

Дата	Продажи, шт.[-2]	Продажи, шт.[-1]	Продажи, шт.	Продажи, шт.[+1]
01.09.2011			45	82
01.10.2011		45	82	120
01.11.2011	45	82	120	192
01.12.2011	82	120	192	229
01.01.2012	120	192	229	161
01.02.2012	192	229	161	

Результирующая таблица при значении *Удалять все неполные записи*:

Дата	Продажи, шт.[-2]	Продажи, шт.[-1]	Продажи, шт.	Продажи, шт.[+1]
01.11.2011	45	82	120	192
01.12.2011	82	120	192	229
01.01.2012	120	192	229	161