

## Содержание

1	Простые алгоритмы и метрики качества	2
2	Загоночная работа	2
3	Лог. КЛШ-2019	3
3.1	Плакат . . . . .	3
4	Решения	3
5	Источники мудрости	3

## Цель

Научить школьников своими руками прогнозировать сезонные временные ряды.

## 1. Простые алгоритмы и метрики качества

месяц	2017	2018	год	население
январь	-	49	2010	143.0
февраль	-	53	2011	142.9
март	-	57	2012	143.1
апрель	-	79	2013	143.4
май	41	40	2014	143.7
июнь	112	103	2015	146.3
июль	150	114	2016	146.5
август	139	150	2017	146.8
сентябрь	126	113	2018	?
октябрь	80	?	2019	?
ноябрь	71	?		
декабрь	73	-		

## 2. Загоночная работа

Кот Матроскин записал надои коровы Мурки за последние четыре месяца: 20, 30, 30, 40. Построй прогноз надоев на один и два шага вперёд с помощью:

1. наивного алгоритма;
2. модели ETS(ANN) с параметрами  $\alpha = 0.5$ ,  $\ell_0 = 30$ ;
3. модели ETS(AAN) с параметрами  $\alpha = 0.5$ ,  $\ell_0 = 20$ ,  $\beta = 0.1$ ,  $b_0 = 6$ .

Полезные уравнения:

$$\begin{cases} y_t = \ell_{t-1} + \varepsilon_t \\ \ell_t = \ell_{t-1} + \alpha \varepsilon_t \\ \hat{y}_{t+h} = \ell_t \end{cases} \quad \begin{cases} y_t = \ell_{t-1} + b_{t-1} + \varepsilon_t \\ \ell_t = \ell_{t-1} + b_{t-1} + \alpha \varepsilon_t \\ b_t = b_{t-1} + \beta \varepsilon_t \\ \hat{y}_{t+h} = \ell_t + h b_t \end{cases}$$

Подсказка: из первого уравнения системы можно выразить  $\varepsilon_t$  и подставить его в остальные :)

### 3. Лог. КЛШ-2019

1. Было 12 школьников. Дал два ряда: годовой и месячный, разбил в группы по двое-трое и предложил спрогнозировать ряд. Затем предложил описать свой алгоритм на бумаге. Затем передать бумажку соседней команде. Затем соседняя команда реализует алгоритм, который ей передали на бумажке. Не ожидал, оказалось ни одного совпадения. Сказал, что возможность репликации — это основа науки. То, что нереплицируемо — не наука. Далее описали простые алгоритмы: прогноз равен среднему, прогноз равен последнему значению, прогноз равен последнему значению плюс последний прирост. Ввели обозначения  $y_t, \hat{y}_{t+h}$ .
2. Резюмировали идеи: реплицируемость. Три самых простых алгоритма: среднее арифметическое, наивный, сезонный наивный. Записали все три формулой. Задались вопросом оценки качества алгоритма. Поделили выборку на две части: обучающую и тестовую. Спрогнозировали с помощью каждого алгоритма наши два ряда (кроме сезонного наивного для годового ряда). Посчитали  $MAE$ .
3. Перешли к рассмотрению несезонного ряда, колеблющегося вокруг медленно меняющейся долгосрочной составляющей. Нарисовал пример ряда, который подходит под эту идею, и примеры, которые не подходят. Осознали на примере одного момента мысль, что на  $l_6$  должны влиять  $l_5$  и  $y_6$ . Качественно,  $l_6$  должен отклоняться от  $l_5$  в ту же сторону, что и  $y_6$ , но слабее. Возьмем, к примеру,  $l_6 = 0.5y_6 + 0.5l_5$ . Взяли другие коэффициенты  $\alpha$ . Написали общую формулу:

$$\begin{cases} \hat{y}_{t+h} = l_t \\ l_t = \alpha y_t + (1 - \alpha)l_{t-1} \end{cases}$$

Взяли пять наблюдений из Nile. Заполнили столбик  $l_t$  и посчитали прогнозы на один и два шага для  $\alpha = 0.5$  и  $l_0 = 900$ . Потом школьники сами повторили результат для  $\alpha = 0.6$  и  $l_0 = 910$ . Многие ленились :) Затем сравнили по  $MAE$ .

4. В экселе на двух более длинных рядах Лена повторила подсчёт с ручным подбором  $\alpha$  и  $l_0$ .
5. Запустили R :) Оценили ETS(ANN) модель для ряда Nile. С делением выборки на две части и оценкой по  $MAE$  функцией `assu`гасу(). Начали аналогичное действие с рядом gas: успели только график построить и поделить выборку на две части.
6. Рассказал про ETS(AAA) модель. С нуля прошли все шаги с рядом gas.

В теховском файле \newpage стоит, чтобы легко было скопировать секцию, для печати двух копий подряд на одном листе. Это позволяет экономить бумагу и время при печати :)

#### 3.1. Плакат

### 4. Решения

### 5. Источники мудрости

передать потом в bib-файл

1. <https://otexts.com/fpp2/>