

Анализ данных 2

Пожилая саламандра

April 2020

1. Реализация решения

1.1. Функции и классы

- Класс `OneStepLinReg` - наш линейный регрессор:
 - `__init__()` - создает экземпляр класса;
 - `fit()` - подбирает весовые коэффициенты;
 - `predict()` - выдаёт предсказанные значения;
 - `coef()` - выдаёт весовые коэффициенты;
 - `display()` - отрисовка графика линейной регрессии.
- `monthly_produced()` - строит графики количества произведенных мечей на основе данных;
- `monthly_broken_metrics()` - вычисляет метрики (п. 2) для сломанных мечей;
- `monthly_stats()` - считает и строит графики вероятностей поломки в определенные месяцы;
- `make_forecast()` - предсказывает вероятность поломки в следующем месяце;

1.2. Библиотеки

- `NumPy`;
- `matplotlib`;
- `pandas`;
- `sklearn.linearmodel`.

2. Результаты и выводы

Будут описаны в презентации.

3. Метрика

Наша метрика зависит от количества сломанных мечей и через сколько времени после изготовления мечей сломалось.

$$metric = \sum_{i \in rows} \frac{Defects}{ReportDate - ProductDate + 1}; \quad (1)$$

$$metric = 100000 / metric; \quad (2)$$

Значение для Harpy.co : 67.01674940043944

Значение для Westeros.inc : 36.09248958259037

Чем больше метрика, тем качественнее мечи у компании.

4. Вклад участников

- *Никита Денисов* - разработка кода
- *Андрей Ловягин* - блокнот .ipynb
- *Михаил Иванков* - README и презентация