Документация к алгоритму оценки нагрузки агента поддержки.

Цель

Необходимо дать оценку уровню нагрузки сотрудника за последнюю неделю (оценка 0 - кол-во выполненных заданий ниже нормы, 1 - норма, 2 - кол-во выполненных заданий выше нормы).

Исходные данные

Датафрейм с полями из таблицы customer_support:

- id. id обращения;
- created. Дата создания обращения;
- updated. Дата обновления обращения;
- status. Статус обращения (solved решено /closed закрыто без решения);
- channel. Канал обращения;
- assignee id. id сотрудника.

Математические концепции

Чтобы дать оценку уровню нагрузки агента за прошедшую неделю, необходимо проследить историю количества решаемых заданий за предыдущие недели. Для этого считается среднее значение (центральное положение) кол-ва выполняемых заданий за неделю в разрезе статуса / статуса и канала. Существует несколько вариантов выбора центрального положения, но я в своем исследовании использовала медиану,



т.к. она наиболее устойчива к выбросам.

Ясно, что агент не может каждую неделю обрабатывать одно и тоже кол-во обращений, поэтому необходимо определить множество кол-ва обращений, которое будет считаться нормой для того или иного сотрудника. Т.к. за центральное положение была взята медиана, то на ее основе определено медианное абсолютное отклонение.

$$MAD = median(|x_i - \tilde{x}|)$$

Доверительный интервал на его основе получается довольно широким, но на результат этот факт не сильно влияет (было проведено сравнение с результатами алгоритма, использующего среднеквадратическое отклонение).

Таким образом, для каждого сотрудника была определена норма кол-ва обрабатываемых обращений за неделю в разрезе статуса / статуса и канала, и посчитана оценка уровня нагрузки. В результате получены два датафрейма со следующими полями:

- assignee id. id сотрудника;
- status. статус;
- count_last_period. Кол-во обработанных обращений за последнее время;
- count mean calc period. Среднее кол-во обрабатываемых обращений;
- count sem calc period. Отклонение от среднего;
- score_value. Оценка уровня нагрузки.

	assignee_id	status	count_last_period	count_mean_calc_period	count_sem_calc_period	score_value
0	4225872478	closed	46	256.5	59.0	0
1	4225872478	solved	0	0.0	0.0	0
0	379332361	closed	53	100.0	7.0	0
0	12604869947	closed	196	188.5	14.5	1
1	12604869947	solved	0	0.0	0.0	0
		577				
1	364622202673	closed	0	0.0	0.0	0
0	365657486013	closed	0	0.0	0.0	0
1	365657486013	solved	0	0.0	0.0	0
0	365658390953	solved	0	0.0	0.0	0

- assignee_id. id сотрудника;
- status. статус;
- channel. канал;
- count_last_period. Кол-во обработанных обращений за последнее время;
- count mean calc period. Среднее кол-во обрабатываемых обращений;
- count sem calc period. Отклонение от среднего;
- score value. Оценка уровня нагрузки.

	assignee_id	status	channel	count_last_period	count_mean_calc_period	count_sem_calc_period	score_value
0	4225872478	closed	help_widget	12	44.5	5.98	0
1	4225872478	closed	facebook	1	0.0	0.73	2
2	4225872478	closed	other	1	0.0	0.50	2
3	4225872478	closed	chat	22	141.5	26.82	0
4	4225872478	closed	call	9	26.5	5.02	0
9	365658390953	closed	help_widget	0	0.0	0.00	0
10	365658390953	closed	zendesk	0	0.0	0.00	0

Далее подсчитан итоговый скоринг (средний score_value) агентов поддержки. Результаты записаны в датафрейм со следующими полями:

- assignee_id. id сотрудника;
- score_value. Оценка уровня нагрузки.

	assignee_id	score_value
0	123193832	0.0
1	288517962	1.0
2	291235568	0.0
3	291458441	1.0
4	291643586	0.0
5	379332361	0.0
6	502786743	0.0
7	774796347	1.0
8	867359386	1.0
9	1162180708	0.0
10	1430684637	0.0

Итог

Произведен анализ данных нагрузки сотрудников поддержки на основе таблицы customer_support. На основе математических концепций удалось выяснить, кто из агентов недогружен/в норме/перегружен.