<https://data.4tu.nl/collections/BPI_Challenge_2014/5065469>

Перевод описания задачи:

Задача состоит в том, чтобы разработать (черновик) прогностическую модель, которую можно использовать для реализации в среде BI.

Целью этой прогностической модели будет помощь компании при внедрении выпусков программного обеспечения для уменьшения нагрузки на службу поддержки и/или ИТ-операции.

Файлы содержат сведения о записи из инструмента ITIL Service Management под названием HP Service Manager. Мы предоставляем вам выписки в формате CSV с номером взаимодействия, инцидента или изменения в качестве идентификатора обращения. Наряду с этими файлами дел мы предоставляем вам журнал действий, связанный с делами об инцидентах.

**BPI Challenge 2014**

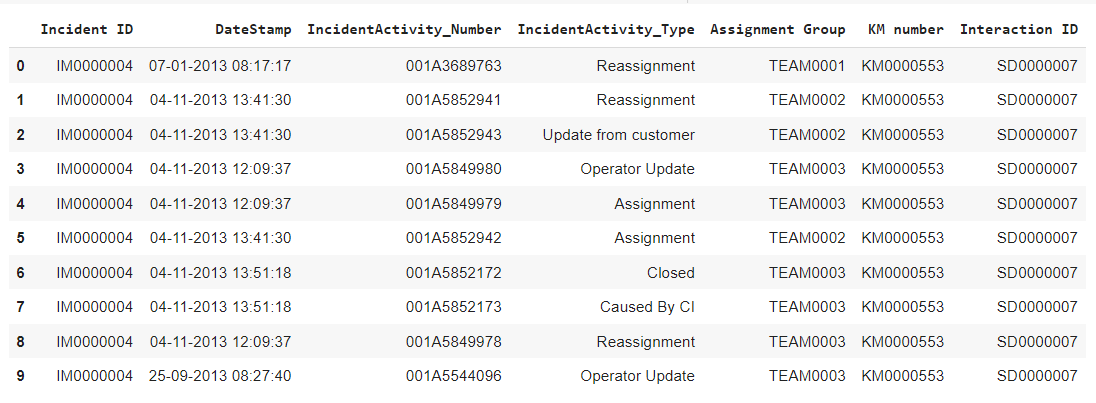
**Описание данных:**

Предоставлено 4 файла в формате csv:

* detail\_incident\_activity - журнал действий для инцидентов
* detail\_incident - записи, относящиеся к инцидентам
* detail\_interaction - записи, относящиеся к взаимодействиям.
* detail\_change - записи, относящиеся к изменениям

Работать будем только с detail\_incident\_activity, поскольку только он содержит журнал событий, в котором 46616 инцидентов.

**Данные в журнале:**



**Преобразование входных данных:**

События «Closed» и «Caused By CI» зачастую происходят в одно время, поэтому сделаем действие «Closed» последним в цепочке событий, чтобы оно всегда было после «Caused By CI». Для этого добавим 1 секунду к времени данного события.

Кроме того, исключим из анализируемых данных инциденты, у которых нет события «Closed», поскольку они еще не завершены и в теории могут исказить собираемую статистику.

**Частотные признаки**

Топ команд по закрытым инцидентам:

|  |  |
| --- | --- |
| TEAM0008 | 81584 |
| TEAM0039 | 19199 |
| TEAM0018 | 17368 |
| TEAM0031 | 16414 |
| TEAM0007 | 15864 |
| TEAM0023 | 15190 |
| TEAM0086 | 13515 |
| TEAM0191 | 11998 |
| TEAM0015 | 11964 |
| TEAM0075 | 11255 |
| TEAM0003 | 10047 |
| TEAM0016 | 9096 |
| TEAM0181 | 8538 |
| TEAM9999 | 8246 |
| TEAM0069 | 7151 |

Топ типов событий:

|  |  |
| --- | --- |
| Assignment | 88247 |
| Operator Update | 56149 |
| Reassignment | 51870 |
| Status Change | 50322 |
| Closed | 50145 |
| Open | 46158 |
| Update | 35945 |
| Caused By CI | 34382 |
| Quality Indicator Fixed | 7791 |
| Communication with customer | 6142 |
| Description Update | 4470 |
| Update from customer | 3900 |
| External Vendor Assignment | 3842 |
| Pending vendor | 3842 |
| Mail to Customer | 3788 |

В нашем журнале событий 31725 различных вариантов путей.

Самые часто встречающиеся цепочки событий:

* ('Open', 'Closed', 'Caused By CI')
* ('Open', 'Status Change', 'Assignment' , 'Caused By CI', 'Closed')

Минимальное и максимальное количество событий в одном случае:

* Минимальное количество событий: 1
* Максимальное количество событий: 178

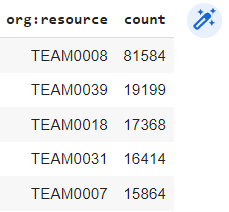
Команда, потратившие больше всего времени:

* Потративший больше всего времени: IM0005897
* Длительность: 103626 days 03:37:45

Команда, потратившие меньше всего времени:

* Потративший меньше всего времени: IM0000006
* Длительность: 0 days 00:00:00

Команды, которые перегружены работой (совершили больше всех действий):



При этом:

* команд 242
* Событий 466737

Таким образом 5 команд делают треть инцидентов.

**Модели**

Для построения моделей данных использовались следующие алгоритмы:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Алгоритм | Простота (Simplicity) | Пригодность воспроизведения (fitness) | Точность (Precision) | Обобщение (generalization) |
| Alpha miner | 0.3889 | 0.2572 | 0.8292 | 0.9288 |
| Dfg graph | 0.3536 | 0.7428 | 1.0 | 0.6561 |
| Inductive miner | 0.5812 | 1.0 | 0.2622 | 0.9162 |
| Heuristics miner | 0.3829 | 0.9946 | 0.36 | 0.8298 |

**Заключение**

В ходе выполнения лабораторной работы был изучен журнал событий из BPI Challenge 2014, получены основные характеристики, сравнены алгоритмы process mining.

Метрики показывают, что Heuristics miner является наиболее подходящим для дальнейшего использования при оптимизации работы компании при внедрении выпусков программного обеспечения и тех поддержке.