Predicting Incident Management Service Level Agreement (SLA) Failures

Executive Summary

Carolyn M. Hennings
Western Governors University
chenn15@wgu.edu
March 10, 2020

Table of Contents

Table of Exhibits

1 Introduction

Information Technology (IT) Service Management practices aim to optimize the efficiency and effectiveness of IT services delivered to users. Incidents cause service disruptions. Service Level Agreements (SLA) establish thresholds for resolution of incidents within specified timeframes based on impact and urgency designations. Decreasing SLA breaches increases the availability of IT services and represents an important consideration for IT service providers.

This study explored indicators of incident SLA breaches with respect to data available during the early stages of an IT incident's lifecycle. The study built a Logistic Regression model using Python and a number of tools from the SciKit-Learn library. Some supplementary analysis leveraged the R language. This overview describes the data collection, the data extraction and preparation, and the analysis steps performed throughout the study followed by a summary of findings and recommendations.

2 Problem Statement and Hypotheses

IT organizations use IT Service Management (ITSM) systems to capture information about the execution of Incident Management processes. These systems produce logs containing details about incidents, for example, the steps taken to resolve them, the individuals involved with the incidents, the elements within the IT environment impacted by the incident, and timestamps for actions taken throughout the lifecycle of an incident.

This project investigated an extract of Incident Management data from an ITSM system to determine indicators of failure to meet an SLA threshold and to develop a model for predicting those incidents. Insight into SLA-at-Risk conditions assists in notifying management of IT environment components requiring attention.

What factors predict Incident Management SLA compliance?

Hypotheses

The hypotheses under study focus on identifying significant factors indicating the probability of an IT support organization's ability to close an incident within agreed service level thresholds. The study leveraged logistic regression techniques to test the following hypotheses:

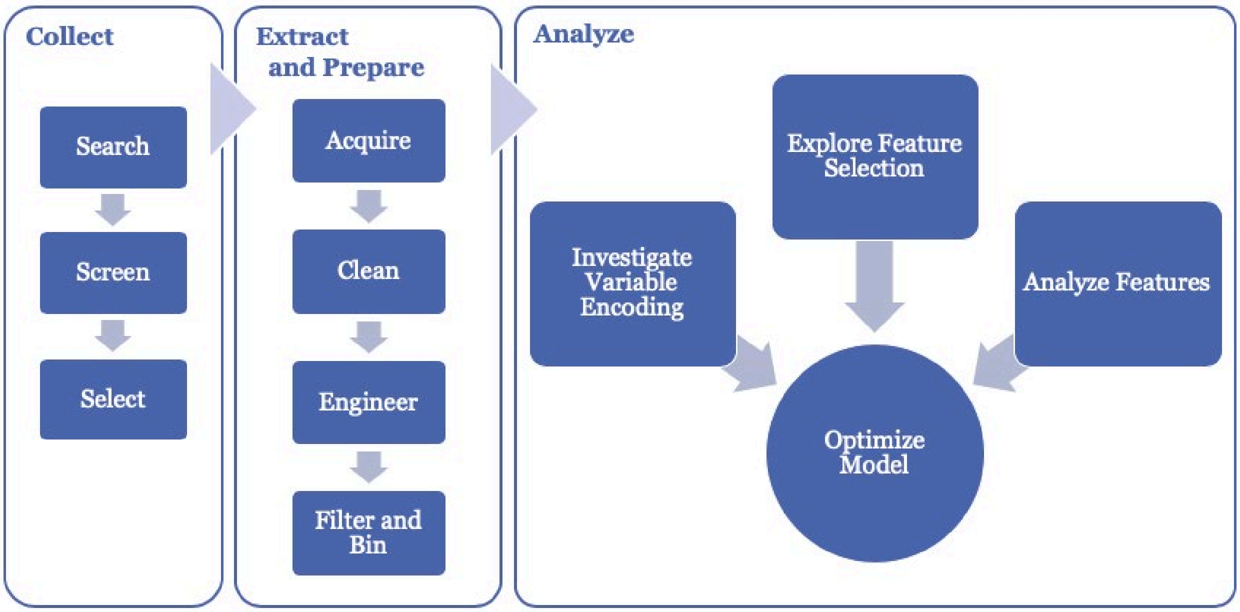
H_0: Data contains no significant indicators about the final SLA status of an incident (the coefficients of a logistic regression model are zero, $\beta_i = 0$)
H_1: Data contains significant indicators about the final SLA status of an incident (the coefficients of a logistic regression model are not zero, $\beta_i \neq 0$)

The factors under consideration were limited to data available in the early stages of an incident's lifecycle. If any one of the logistic regression model's coefficients significantly differs from zero, the study will accept the alternative hypothesis (H_1), otherwise the study will fail to reject the null hypothesis (H_0).

3 Data Analysis Process

The study consisted of three phases consisting of multiple steps as shown in Exhibit 1.

Exhibit 1. Data Analysis Process



3.1 Collect

A study investigating factors contributing to incident management SLA risk requires an extract from an ITSM system used by an IT organization for tracking incidents over a specific period. An internet search identified and screened multiple applicable data sources. This study leveraged an existing, publicly-available data set used in the *2014 Business Processing Intelligence Challenge (BPIC)* (*10th International Workshop on Business Process Intelligence 2014*, n.d.). While the challenge released four data sets, this project focused only on the Incident Records file (Van Dongen, 2014). The selected data set consists of 46,606 observations having 28 variables extracted from an ITSM system used by a bank located in the Netherlands (*Quick reference BPI Challenge 2014*, n.d.). The terms for use of the data set specify that "The user is allowed to remix, transform or build upon the data, but only for noncommercial purposes" (4TU.Centre for Research Data, 2016).

3.2 Extract and Prepare

Acquire. The data set required downloading from the *2014 Business Processing Intelligence Challenge (BPIC)* website located at https://www.win.tue.nl/bpi/doku.php?id=2014:challenge. Initial exploratory data analysis (EDA) identified a relatively clean data set. Preliminary data cleaning steps included: conversion of strings representing dates to datetime data type, removal of non-incident records, and removal of records with a status other than `closed`.

Clean. Data profiling during the previous step revealed collinearity among some variables. This step addressed some collinearity through creating aggregated variables and noted other items for later consideration. The project addressed missing values by dropping records representing fewer than 4% of the total, setting values to zero, "Not Applicable", and "Yes/No" where appropriate (Nisbet et al., 2009, pp. 50–75).

Engineer. The original data source lacked a binary indicator for the target variable. The project engineered the target variable, `SLAFail`, by setting the value to `1`, according to the business rules described in Exhibit 2.

Exhibit 2. Business Rule for SLAFail Target Variable

Priority	Time Between Opened and Resolved
1 Very High	Greater than 240 minutes (4 hours)
2 High	Greater than 480 minutes (8 hours)
3 Medium	Greater than 1440 minutes (1 day)
4 Low	Greater than 2880 minutes (2 days)
5 Very Low	Greater than 5760 minutes (4 days)

Filter and Bin. This final data extraction and preparation step addressed the high dimensionality of datetime variables, collinearity, and restricted the data set based on timeframe.

3.3 Analyze

Analysis developed a logistic regression model that predicts the status of `SLAFail`. Analysis investigated three feature-focused aspects: variable encoding, feature selection, and feature analysis. Optimization followed an iterative approach to refining the model. Throughout the analysis process, decisions stemmed from review of classification accuracy rates and the Area Under the Curve (AUC) score obtained from the Receiver Operating Characteristic (ROC) curve diagnostic (Tufféry, 2011, pp. 454, 458).

The optimization step leveraged the information and insight gained from the previous three steps to develop a logistic regression model that predicts SLA compliance at the early stages of an incident's lifecycle. With a Null Accuracy of 70%, optimization efforts aim at achieving a greater Classification Accuracy score. Given the inability to obtain more training data or to add features, this analyst's optimization choices were limited to searching for less complicated/flexible models and looking at more complicate/flexible models (Vanderplas, n.d.).

4 Findings

In response to the problem statement, this section summarizes analysis implications in the context of early identification of incidents likely to cause breaches in SLA thresholds.

What factors predict Incident Management SLA compliance?

Given the study's problem statement and hypotheses, the project developed a logistic regression model predicting the final SLA status of an incident based on data available during the early stages of an incident's lifecycle. Recall:

H_0: Data contains no significant indicators about the final SLA status of an incident (the coefficients of a logistic regression model are zero, $\beta_i = 0$)
H_1: Data contains significant indicators about the final SLA status of an incident (the coefficients of a logistic regression model are not zero, $\beta_i \neq 0$)

If any one of the logistic regression model's coefficients significantly differs from zero, the study will accept the alternative hypothesis (H_1), otherwise the study will fail to reject the null hypothesis (H_0). As shown in Exhibit 3, seven of the eight variables showed a significance level (P-value) less than alpha (0.05). The study accepted the alternative hypothesis that data does contain significant indicators of an incident's final SLA status.

Exhibit 3. Logistic Regression Formula Coefficients

Variable	Coefficient	F score	P-value
Service_Component_WBS_aff	0.1040	141.5840	0.0000
Urgency	0.1881	6104.4400	0.0000
KM_number	1.1217	177.7560	0.0000
Count_Related_Interactions	1.1726	679.7630	0.0000
Count_Related_Incidents	1.2459	902.1090	0.0000
Count_Related_Changes	0.0388	62.0758	0.0000
Open_Time_HourOfDay	0.1824	5.6072	0.0179
Open_Time_DayOfWeek	-0.0358	0.3519	0.5530

While the study provided statistical evidence for the acceptance of the alternative hypothesis, multiple limitations restrict the practical value of the resulting model. The best model identified by the study achieved a classification accuracy score of 78%, while the null accuracy (always selecting the majority event) resided at 70%. Given implementation costs, this analyst questions the value gained with only an 8% accuracy increase. Also, given the nature of the source data set containing anonymous information, the study was unable to investigate additional factors that could provide further discrimination and additional benefit to the model. Categories within Knowledge Management articles, as well as data representing configuration item types, could potentially improve the model.

5 Limitations of Techniques and Tools

The study leveraged a number of statistical and data analysis techniques, each with benefits and limitations. Feature selection and factor analysis techniques contributed to identifying the factors contributing the most to SLA failures. One limitation of these techniques, along with the application of a logistic regression model, included the need to manipulate categorical data into numeric values. Additional limitations stemmed from the selection of Python as a development environment. While a cost effective and comprehensive solution, this analyst's computing environment experienced resource constraints while attempting to perform factor analysis and needed to leverage R as an alternative.

6 Proposed Actions

Based on these results, this analyst recommends a course of action focused on further investigation of the specific Configuration Items and Knowledge Management articles indicating incidents that exceed SLA thresholds. Analysis also showed that a subset of variables provides Incident Management professionals with direction for swift identification of incidents that may breach an SLA threshold. Principal component analysis and the hierarchical cluster analysis both grouped the following variables together:

- `Service_Component_WBS_aff`
- `KM_Number`
- `CI_TypeSubType_aff`

With this understanding, coupled with the identification of `Service_Component_WBS_aff` and `KM_Number` as significant indicators of SLA-at-Risk, Incident Management professionals can prioritize attention on identifying specific items in the IT environment causing incidents that exceed SLA thresholds.

7 Benefits

While the study formally accepted the Alternative Hypothesis, the marginal gain in predictive capability (from 70% to 78%) raises a question about the value of implementing the model in a production environment. However, the identification of Configuration Items and Knowledge Articles as significant indicators contribute to benefits available to customers and number of roles within an IT support organization. With the information provided by the study, Service Desk agents can swiftly identify incidents likely to breach an SLA and can prioritize resolution efforts over other activities. Incident Managers, those responsible for the overall Incident Management process, gain the ability to reassign resources for focus on critical incidents with a long-term benefit of overall reduction in the duration of incidents. These actions will increase service availability for users and subsequently increase customer satisfaction.

8 References

4TU.Centre for Research Data. (2016, August). *General terms of use for 4TU.Centre for Research Data*. https://data.4tu.nl/repository/resource:terms_of_use/PDF

10th International Workshop on Business Process Intelligence 2014. (n.d.). https://www.win.tue.nl/bpi/doku.php?id=2014:challenge

Nisbet, R., Elder, J. F., & Miner, G. (2009). *Handbook of statistical analysis and data mining applications*. Academic Press/Elsevier.

Quick reference BPI Challenge 2014. (n.d.). Retrieved February 17, 2020, from https://www.win.tue.nl/bpi/lib/exe/fetch.php?media=2014:quick_reference_bpi_challenge_2014.pdf

Tufféry, S. (2011). *Data Mining and Statistics for Decision Making: Tufféry/Data Mining and Statistics for Decision Making*. John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470979174

Van Dongen, B. F. (Boudewijn). (2014). *BPI Challenge 2014: Incident details*. Rabobank Nederland. https://doi.org/10.4121/UUID:3CFA2260-F5C5-44BE-AFE1-B70D35288D6D

Vanderplas, J. (n.d.). *Hyperparameters and Model Validation*. Python Data Science Handbook. Retrieved March 7, 2020, from https://jakevdp.github.io/PythonDataScienceHandbook/05.03-hyperparameters-and-model-validation.html