**Prediction of target SLAs are met within businesses.**

**Introduction**

Business context, stakeholders and value

Service Level Agreements (SLAs) define the level of service a customer can expect from a seller or marketplace. One of the most important service clauses in an SLA is an **assurance that the seller will respond to a customer query within a given time period**. <https://support.edesk.com/setting-sla-targets>

**Most businesses profitability relies on Target SLAs**. Knowing the fact if the job is possible to complete within a certain time frame or if not, discovering drawbacks for failures assists a business to prosper better. Making data driven decisions affects all identified five core business systems such as,

* Sales & Marketing.
* Quality & Product/Service Delivery.
* Product Development.
* Accounting & Technology.
* Administrative (Management, HR & Finance)

Below report is an example of checking if the existing system is doing its part for the business. **It also allows to highlight where the system needs further data to have a complete picture of the business**. All target driven customer-oriented businesses do benefit immensely by predicting a timeframe for certain jobs. It is one of the major turning points for customers to decide if the job is worth going ahead or not by assessing the profitability, affordability, outcome of the end result and many more.

‘Incident management process enriched event log’ is used as an example for further explanation of this concept.

**Incident management process enriched event log**

**Abstract**: This event log was extracted from data gathered from the audit system of an instance of the ServiceNow platform used by an IT company. It is enriched with data loaded from a relational database underlying a corresponding process-aware information system. Information was anonymized for privacy.

<https://archive.ics.uci.edu/ml/datasets/Incident+management+process+enriched+event+log>

Number of instances: 141,712 events (24,918 incidents)  
Number of attributes: 36 attributes (1 case identifier, 1 state identifier, 32 descriptive attributes, 2 dependent variables)  
  
The attribute ‘closed’ is used to determine the dependent variable for the time completion prediction task. The attribute ‘resolved’ is highly correlated with ‘closed’. In this event log, some rows may have the same values (they are equal) since not all attributes involved in the real-world process are present in the log.  
  
Attributes used to record textual information are not placed in this log.  
  
Missing values should be considered ‘unknown information’.

**Attribute Information:**

1. number: incident identifier (24,918 different values);  
2. incident state: eight levels controlling the incident management process transitions from opening until closing the case;  
3. active: boolean attribute that shows whether the record is active or closed/cancelled;  
4. reassignment\_count: number of times the incident has the group or the support analysts changed;  
5. reopen\_count: number of times the incident resolution was rejected by the caller;  
6. sys\_mod\_count: number of incident updates until that moment;  
7. made\_sla: boolean attribute that shows whether the incident exceeded the target SLA;  
8. caller\_id: identifier of the user affected;  
9. opened\_by: identifier of the user who reported the incident;  
10. opened\_at: incident user opening date and time;  
11. sys\_created\_by: identifier of the user who registered the incident;  
12. sys\_created\_at: incident system creation date and time;  
13. sys\_updated\_by: identifier of the user who updated the incident and generated the current log record;  
14. sys\_updated\_at: incident system update date and time;  
15. contact\_type: categorical attribute that shows by what means the incident was reported;  
16. location: identifier of the location of the place affected;  
17. category: first-level description of the affected service;  
18. subcategory: second-level description of the affected service (related to the first level description, i.e., to category);  
19. u\_symptom: description of the user perception about service availability;  
20. cmdb\_ci: (confirmation item) identifier used to report the affected item (not mandatory);  
21. impact: description of the impact caused by the incident (values: 1 High; 2 Medium; 3 Low);  
22. urgency: description of the urgency informed by the user for the incident resolution (values: (values: 1 High; 2 Medium; 3 Low);  
23. priority: calculated by the system based on 'impact' and 'urgency';  
24. assignment\_group: identifier of the support group in charge of the incident;  
25. assigned\_to: identifier of the user in charge of the incident;  
26. knowledge: boolean attribute that shows whether a knowledge base document was used to resolve the incident;  
27. u\_priority\_confirmation: boolean attribute that shows whether the priority field has been double-checked;  
28. notify: categorical attribute that shows whether notifications were generated for the incident;  
29. problem\_id: identifier of the problem associated with the incident;  
30. rfc: (request for change) identifier of the change request associated with the incident;  
31. vendor: identifier of the vendor in charge of the incident;  
32. caused\_by: identifier of the RFC responsible by the incident;  
33. close\_code: identifier of the resolution of the incident;  
34. resolved\_by: identifier of the user who resolved the incident;  
35. resolved\_at: incident user resolution date and time (dependent variable);  
36. closed\_at: incident user close date and time (dependent variable).

Explore the dataset for possible prediction.

When a job comes through knowing if it can be completed within target SLAs are extremely important. In order to assess, all related data needs to be identified and manipulated to find a base result so that predictions can be made.

Answering the question ‘**What affects a new case?**’ allows to identify all attributes related to a new case.

* opened at: incident user opening date and time;
* contact type: categorical attribute that shows by what means the incident was reported;
* location: identifier of the location of the place affected;
* category: first-level description of the affected service;
* subcategory: second-level description of the affected service (related to the first level description, i.e., to category);
* u\_symptom: description of the user perception about service availability;
* impact: description of the impact caused by the incident (values: 1 High; 2 Medium; 3 Low);
* urgency : description of the urgency informed by the user for the incident resolution (values: 1 High; 2 Medium; 3 Low);

Out of given data, predicting the possibility of closing a case within the target time frame is possible.

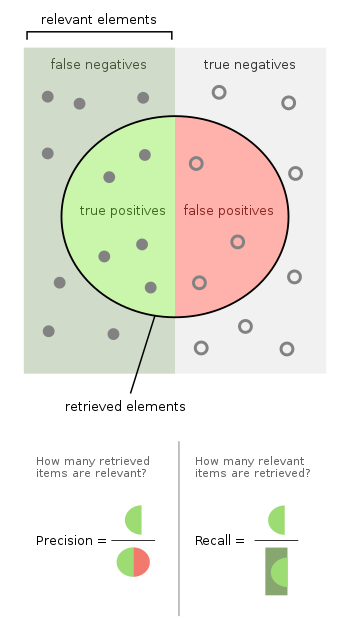
possible prediction variables

* Time to complete
* made\_sla: boolean attribute that shows whether the incident exceeded the target SLA;

Assumption: SLA is same for all alerts.

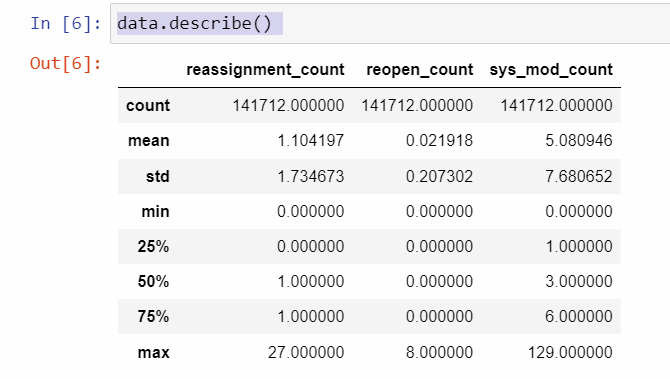
Incident exceeded the target SLA means that the target time got exceeded. Therefore, False is what is expected. That is false positive. So **false negative needs to be low in this case**.

|  |  |  |
| --- | --- | --- |
| incident within the target SLA | identified as exceeded | True Positive |
| incident within the target SLA | not identified as exceeded | True Negative |
| incident exceeded the target SLA | identified as exceeded | False Positive |
| incident exceeded the target SLA | not identified as exceeded | False negative |



Summery of the data frame is given below.

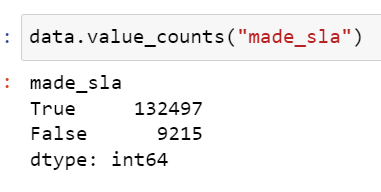
* reassignment\_count: number of times the incident has the group or the support analysts changed;
* reopen\_count: number of times the incident resolution was rejected by the caller;
* sys\_mod\_count: number of incident updates until that moment;



By default, describe() computes and only shows statistics for the numeric variables.

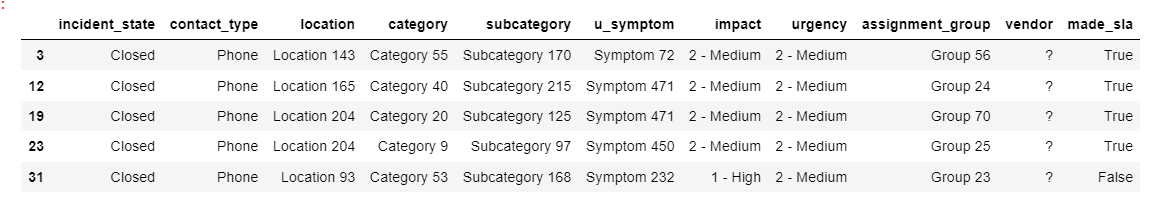
* count
* mean
* standard deviation
* the minimum and maximum
* the 25th, 50th, and 75th percentiles

Count of jobs ‘did get completed within SLA’ vs ‘did not get completed within SLA’.

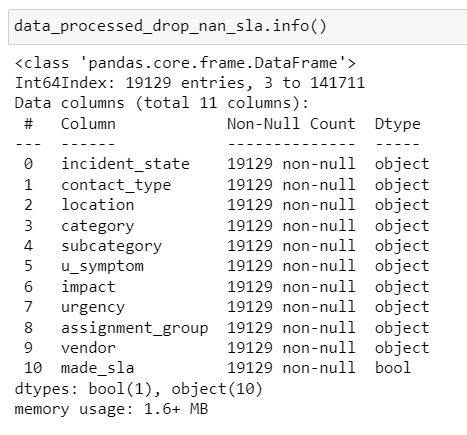


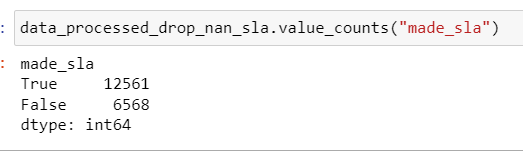
To have a clear idea of the dataset, it is necessary to drop unwanted columns. Therefore, all unusable columns which are not useful at the time of a new job were dropped.

### Selecting all columns that is known as a new job along with the prediction variable.



Exploring the filtered dataset





Although data is lost due to filtering, it is ignorable.

|  |  |
| --- | --- |
| Before | After |
|  |  |

Location, Category and Group VS value counts

|  |  |
| --- | --- |
|  | There are six locations clearly indicating that it is more likely to generate a new case. |
|  | Categories 46, 53, 26, and 42 seems to be more popular compared to rest of the categories. |
|  | This graph clearly indicates that Group 70 is the busiest group out of all.  It also shows that a big percentage of group information is not entered. |

Above graphs clearly shows where and which categories the business flourishes. Assessing that, groups(staff) can be allocated to satisfy the need. Further hiring and reallocation of staff also can be done with a better understanding of this aspect of the business.

# [**https://towardsdatascience.com/accuracy-recall-precision-f-score-specificity-which-to-optimize-on-867d3f11124**](https://towardsdatascience.com/accuracy-recall-precision-f-score-specificity-which-to-optimize-on-867d3f11124)

# **Which performance metric to choose?**

## **Accuracy**

It’s the ratio of the correctly labeled subjects to the whole pool of subjects.  
Accuracy is the most intuitive one.  
**Accuracy answers the following question: How many students did we correctly label out of all the students?**  
**Accuracy = (TP+TN)/(TP+FP+FN+TN)**numerator: all correctly labelled subject (All trues)  
denominator: all subjects

## **Precision**

Precision is the ratio of the **correctly**predicted out of all predicted values.  
**Precision answers the following: How many of those who we predicted are actually true.**

**Precision = TP/(TP+FP)**

## **Recall (aka Sensitivity)**

Recall is the ratio of correctly predicted value.

**Recall answers the following question: Of all the people who are diabetic, how many of those we correctly predict?  
Recall = TP/(TP+FN)**numerator: +ve labeled diabetic people.  
denominator: all people who are diabetic (whether detected by our program or not)

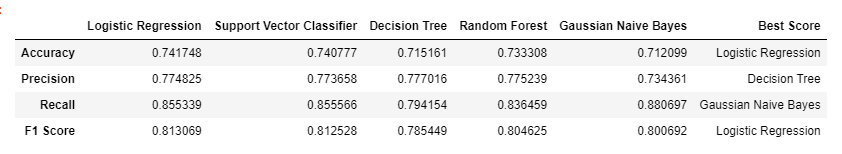
## **F1-score (aka F-Score / F-Measure)**

F1 Score considers both precision and recall.  
**It is the harmonic mean(average) of the precision and recall.**. Oppositely F1 Score isn’t so high if one measure is improved at the expense of the other.  
For example, if P is 1 & R is 0, F1 score is 0.  
**F1 Score = 2\*(Recall \* Precision) / (Recall + Precision)**

CatBoost Classifier

|  |  |
| --- | --- |
|  |  |

Comparing other possible classifiers.



Confusion Matrix for other classifiers.

|  |  |
| --- | --- |
| SVM | Random Forest Model |
|  |  |
| Decision Tree Classifier | CatBoost Classifier |
|  |  |

Above comparison indicates that catboost and RandomForrest model seems to be doing better predictions. Both classifiers got lower False Negative values.

Random Forest Model (Gaussian Classifier)

|  |  |  |
| --- | --- | --- |
| n\_estimators=90 | n\_estimators=100 | n\_estimators=80 |
|  |  |  |
|  |  |  |

Changing n\_estimators Random Forest Model was tested. Out of all, n\_estimators=90 did the best of all three as it provided 9.48% for FN. It also gave a relatively higher Accuracy score.

Comparison of Catboost with different parameters.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| CatBoost | 'depth': 8,  'iterations': 600,  'learning\_rate': 0.05, | 'depth': 8,  'iterations': 600,  'learning\_rate': 0.1, | 'depth': 10,  'iterations': 600,  'learning\_rate': 0.05, | 'depth': 12, 'iterations': 600, 'learning\_rate': 0.1, |
| accuracy\_score | 0.7628506708485799 | 0.759539989545217 | 0.7612824533890922 | 0.7576232793169542 |
| precision\_score | 0.7970871389780301 | 0.7953983176645225 | 0.7925817471937531 | 0.7909424724602203 |
| recall\_score | 0.8569532908704883 | 0.8532377919320594 | 0.861995753715499 | 0.857484076433121 |
| f1\_score | 0.8259368205652897 | 0.823303457106274 | 0.8258326976862446 | 0.8228702406723546 |
| Confusion Matrix |  |  |  |  |
| FN of Matrix | 9.39% | 9.64% | 9.06% | 9.36% |

After comparing different parameters, three out of four indicated that catboost does better compared to Random Forest Model as most FN percentages indicated lower values than produced 9.48% for FN.

Out of trained CatBoost Classifier parameters, 'depth': 10, 'iterations': 600, 'learning\_rate': 0.05, setting did return 9.06% of a FN percentage. It also produced the highest Accuracy\_Score of 0.7612824533890922.

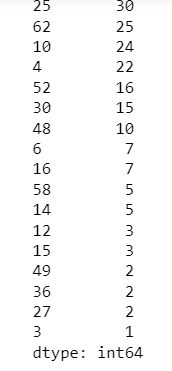
Therefore, it is certain that **Catboost Classifier is the most suitable for this dataset**.

Going further assessing the data set, it seems to be majority of details are missing.

If these columns were populated, there can be a comparison of Vendor, Assignment group, Location and Category.

|  |  |  |
| --- | --- | --- |
|  |  |  |

**Category Column got unnecessary categories** that can slow down the process. It also can be avoided by categorising several main categories and then sub categories where many can handle instead of limited personals. There is also a possibility of completely removing those categories after checking the impact on the data set and the business.



Comparing the location and the assignment group, it is possible to predict the possibility of needing certain resources. It will also allow the business to reduce unnecessary costs by re allocating staff and resources to much need areas of business.

Comparing the groups with target SLAs, it is possible to eliminate unnecessary costs of the business as well as identify and compliment groups who are meeting targets well and beyond target SLAs.

Conclusion.

Knowing the Target SLAs can be met and what can be further done to reach a business’s optimal goal can be predicted by ensuring certain data is recorded. Given example dataset indicated that CatBoost is the best training classifier to predict its Target SLAs. Recording and training data and identifying further needed data can optimise most customer-oriented businesses to their best.