

# Technical Disclosure Commons

---

Defensive Publications Series

---

April 02, 2019

## REQUISITION TIME TO CLOSE PREDICTOR BASED ON HISTORICAL DATA USING SVM

HP INC

Follow this and additional works at: [https://www.tdcommons.org/dpubs\\_series](https://www.tdcommons.org/dpubs_series)

---

### Recommended Citation

INC, HP, "REQUISITION TIME TO CLOSE PREDICTOR BASED ON HISTORICAL DATA USING SVM", Technical Disclosure Commons, (April 02, 2019)  
[https://www.tdcommons.org/dpubs\\_series/2106](https://www.tdcommons.org/dpubs_series/2106)



This work is licensed under a [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).

This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.

# Requisition Time to Close predictor based on historical data using SVM

## Abstract

*One of the key activities in any Human Resources organization is recruiting the best professionals for the teams they support.*

*One of the issues faced is the apparent unpredictability of the timings during the recruitment process, in other words the number of days it will take to have a new professional working after the moment the vacancy is generated. This leads to difficulties in accurately planning the budgets, workforce and time distribution, and resources allocation.*

*To give more clarity on this process and allow a better business planning, we have developed an algorithm that based on a set of features of the vacancy will help the hiring manager or HR professional predict with improved accuracy when a vacancy will be filled.*

## Problem statement

Until now the main approach to predict the time to close a vacancy, was the following:

Assume it is very difficult to predict how many days it needed to fill a vacancy, and therefore open it in the same moment or as close to the date the budget and planning allowed to do so. This solution is very inefficient, leading to a discrepancy between the date of budget and resources allocation and the date the vacancy is filled to be equal or very close the days it took for a requisition to be filled, many times being several months.

## Our solution

To generate the tool, we need a set of observations of previously filled vacancies. Each vacancy will need to have the following set of features assigned:

- Time to Close: Continuous variable, consisting in number of days it took to have the vacancy filled since it was originally created.
- Geographical Region: Categorical nominal variable, consisting in the geographical region in which the vacancy was filled.
- Business: Categorical nominal variable, consisting in the organization or business the vacancy is placed.
- Job function: Categorical nominal variable, based on the job functions the role has assigned.
- Hierarchy Level: Categorical ordinal variable, based on the position the vacancy is placed against the company's hierarchy, typically dependent on the level of responsibility and accountability or the years of experience and previous roles the candidate is expected to have done in the past.
- Purpose: Categorical nominal variable which will explain the reason the vacancy was published, for example due to attrition, to an strategical investment, or other reasons.
- Instruction: Categorical nominal variable, indicates if the position is open to external applicants (candidates who are not working in the company at the moment), or internal applicants.

- Employee type: Categorical nominal variable which describes if the position is either permanent or temporary.

Even if we have more features describing the filled vacancies in the past, it is important that we use the features we can know before having a selected candidate. Therefore, it is critical that the features describe the vacancy and not the candidate.

Depending on the volume of data and the characteristics of it we should adjust accordingly the granularity of the same. For example, if we are mainly filling vacancies in Spain, the Geographical Region features can be cities. If our vacancies are very spread around Europe, we would use countries instead. It has been proven that with this refinement small improvements can be achieved.

Once the data is available and queryable, the algorithm follows this process:

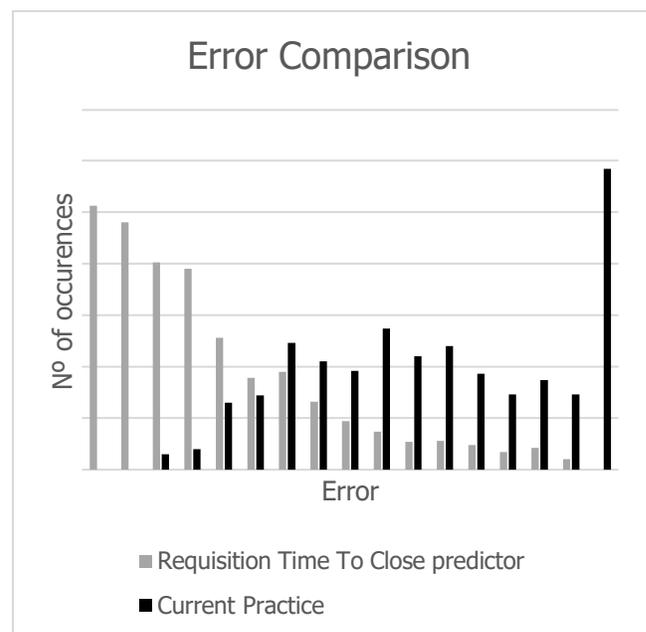
1. Remove incoherent measures and outliers.
  - 1.1. Remove measures using domain specific rules. For example, doesn't make sense that a vacancy has been filled in 1 day or in 1 year.
  - 1.2. Remove outliers using a gaussian threshold of  $3 \times \sigma$ .
2. Split data set into training and test sets using a proportion of 80/20.
3. Train a regression SVM model with the training set. The input parameters are optimized using using a subset of the training set and a grid optimizer algorithm.
4. The model is validated using the test set.
5. The model is persisted into disk for later use.

## Evidence the solution works

Applying the Requisition Time to Close model we will be able to provide an approximation of how many days will a vacancy be open.

On the right we show the result of using the model in 1481 vacancies: clearly the error, measured as the absolute difference between the actual time to close a requisition and the predicted time, when following the standard practice of opening the vacancy in the same period of time we planned the vacancy to be filled (black) is very shifted to the right (higher error), while the prediction given by the Requisition Time to Close model (grey) is very shifted to the left, with the largest group of vacancies having an error of or very close to 0.

The time to close a vacancy, even if aggregated data, is being considered confidential by the Human Resources department, therefore we are unable to provide the actual values in the graph.



In this exercise the model accuracy gives an improvement of 65% less error than not using it. 33% of the vacancies

had an error lower than 19 days, while only a 2% had an error lower than 39 days when not using the model.

With Requisition Time To Close predictor we can have two main advantages:

- Time to Close: If we know the date a vacancy is “opened”, we will know which day to expect this vacancy to be closed with a successful candidate. For example: we had some unexpected attrition, we can plan for how many days the rest of the workforce will need to cope with the cap the attrition generated.
- When to Open: If we know when we actually need a vacancy to be closed, we will know which day we need to open it. For example: we need to start a project in 4 months’ time. Knowing how many days it will take to incorporate a candidate, we can invest the difference between the 4 months and days to fill the vacancy in other recruitment processes.

Some positive repercussions on the previous advantages are:

- Budget allocation: An important part of most organization’s budget is human resources (mainly salary among other related expenses). If we are planning for a vacancy to be filled at a certain date, and we end up having this person start one or two month later, it means that we are not using that piece of budget efficiently in other areas.
- Workforce Planning: Having a vacancy opened longer than the target start date for the new employee, means there are certain number of tasks or assignments which were expecting to be done by someone but weren’t actually done. This can imply having other employees taking care of the un-attended responsibilities, the expected new hire not starting to learn the new role and therefore delaying the deliverables of the team, etc.
- Recruitment efforts efficiency: generally, any Human Resources department will have a taskforce dedicated to recruitment actions to fill vacancies. Having a better understanding on timings will help them prioritize their recruitment efforts, as well as managing better the expectations of the stakeholders such as the hiring managers.

## Competitive approaches

We currently have no evidence of other solutions available in the market and used by our competitors in a wide scale.

## Current status

The Requisition Time to Close predictor based on historical data using machine learning has been submitted for patent disclosure on February 2019 and is currently under revision. It has been implemented using Python linked to an Excel user interface using VBA. We are working on creating a Sharepoint based application in order to make it available for the Human Resources and Hiring Managers community in an easy and efficient way.

## Next steps

The next step would be applying machine learning for other business cases which might have a big impact on HP as a company – and from our point of view, by definition anything related to People has a big impact on the company, both in terms of human capital and finances.

We are also working in improving the model accuracy by using other algorithms as predictors, mainly focusing on Regression Trees and small Neural Networks.

*Disclosed by Aleix Muntal and Quim Muntal, HP Inc.*