



Data Science Case Study

Surgeon OR Scheduling Optimization

Mosaic helped a leading hospital with surgeon scheduling optimization of elective surgery and created a better daily rhythm for surgeons.



Industry

Healthcare



Use Case

Optimizing Surgeon Scheduling



Techniques

Gradient Boosting, Optimization, Usability



Outcome

78% accuracy increase in surgery time predictions

INTRODUCTION

A leading hospital system in the Midwest contacted Mosaic to help with fortifying scheduling decisions as part of a wider transformation effort. Our ultimate objective was to optimize the scheduling of elective surgery and create a better daily rhythm for surgeons. Surgeons have critical, complex jobs. The more support hospitals can provide for these skilled workers, the better outcomes they can hope to achieve, mitigating litigious risk and improving patient health while handling more cases.

Optimizing surgery schedules and providing predictive insights around a complex procedure such as surgery seems like a natural place for hospitals to refine operations. Yet, Mosaic found the industry standard to be ripe for improvement. In the following case study, Mosaic was asked to build a prediction model that could estimate the duration, in minutes, of a given surgery.

THE PROBLEM

Accurately predicting surgery duration has been historically difficult in part due to the availability of data. [Most studies](#)¹ only have access to around 50,000 cases over multiple years, and specific surgery types are not daily occurrences. Additionally, many of the variables associated with a case running longer than expected are unknown until after the procedure has begun. The lack of an accurate prediction can cause significant scheduling problems, depending on the direction of the error. Apparent issues include underutilized operating rooms in the case of overestimation and unplanned overtime in underestimation.

In this case, the principal concern is avoiding significant gaps in a surgeon's OR schedule. These gaps can cause a busy surgeon to leave the OR for a clinic appointment or an appointment at another hospital altogether. When this happens, it can cause disruptions in the rhythm of the OR for everyone. For example, a case might finish early, but the next one can't start right away because the surgeon is no longer in the OR. More accurate predictions help determine a more realistic time required to get a surgery done, helping Mosaic's data scientists in the optimization phase to determine the appropriate block size on the calendar.

CURRENT APPROACH

Upon inspection of the data, it was immediately apparent that the current industry standard for estimating duration called for significant improvement. While the exact method can vary between hospitals, the industry standard typically takes an average of the last ten times the surgeon performed the procedure. While this method is simple to calculate, Mosaic's data scientists found the method overestimating the actual duration by 19% with an average absolute error of 29 minutes. The method was being biased by outliers—surgeries are more likely to take longer than expected than shorter. If a typical case takes an hour, the chances of it taking twice as long (due to complications) are higher than the chances it will take only half the time, for example.



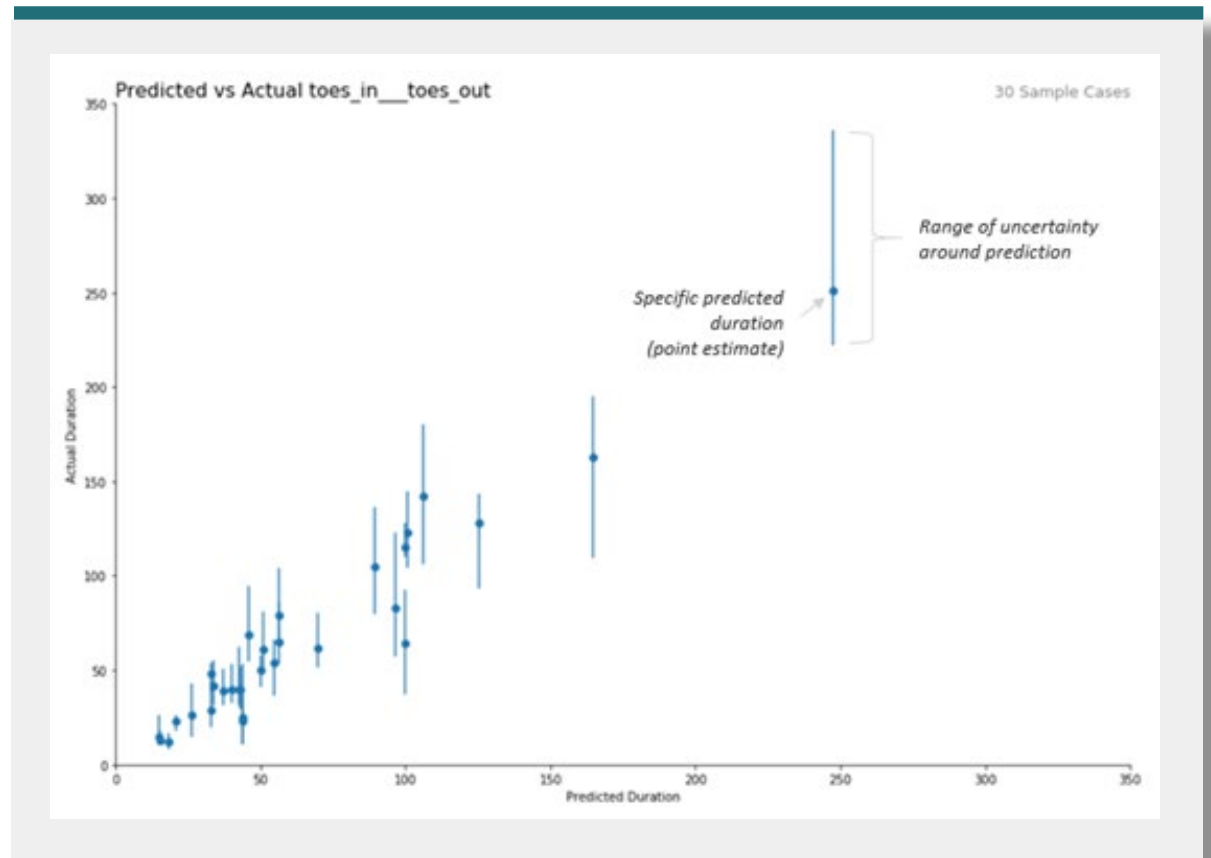
PREDICTIVE MODELING

While many unknown variables can arise once a procedure has begun, several variables known beforehand help inform a better estimate. Known attributes include the type of surgery (elective, urgent, emergency), patient attributes (age, sex, BMI, known medical problems, allergies, etc.), the anesthesiologist, the staff, equipment requirements, etc. Mosaic applied a feature importance & selection algorithm to decide which variables to include in the predictive model with such a broad set of factors. All of the attributes from above, including the procedure, surgeon, and default estimate, were used in the final set of models developed for the client.

The final models predicted *toes-in* to *toes-out* duration (total time the patient occupies the OR) as well as the time required to clean up and set up the room for the next case. Mosaic tested seven different algorithms including, lasso regression, ridge regression, and random forest, to name a few candidates. Gradient boosting regression was ultimately used with a quantile loss function to provide the ability to generate a point-estimate and

a prediction interval of estimated total duration. Mosaic computed over 150 model experiments before settling on the final model.

Mosaic built this model with an eye towards an optimization phase. Predicting time intervals is an excellent insight, but if schedulers don't use this to change decisions, the true power of prescriptive analytics will be lost in the effort. The interval prediction is essential to the optimization phase as the decision to schedule can be seen as a function of the risk tolerance of under or overestimating the actual duration. In other words, schedulers wouldn't necessarily want to use the best prediction if administrators need to be more conservative or risk-tolerant with a given scheduling action.



GRADIENT BOOSTING RESULTS

The final *toes-in* to *toes-out* model resulted in a 38% reduction in MAPE and MAE (average absolute error in terms of percentage and minutes, respectively), and an 85% reduction in bias (significant overestimation to minimal overestimation). An additional metric tracked the prediction “win rate” when comparing the absolute error of the final model with the default estimation method (within a 5-minute window).

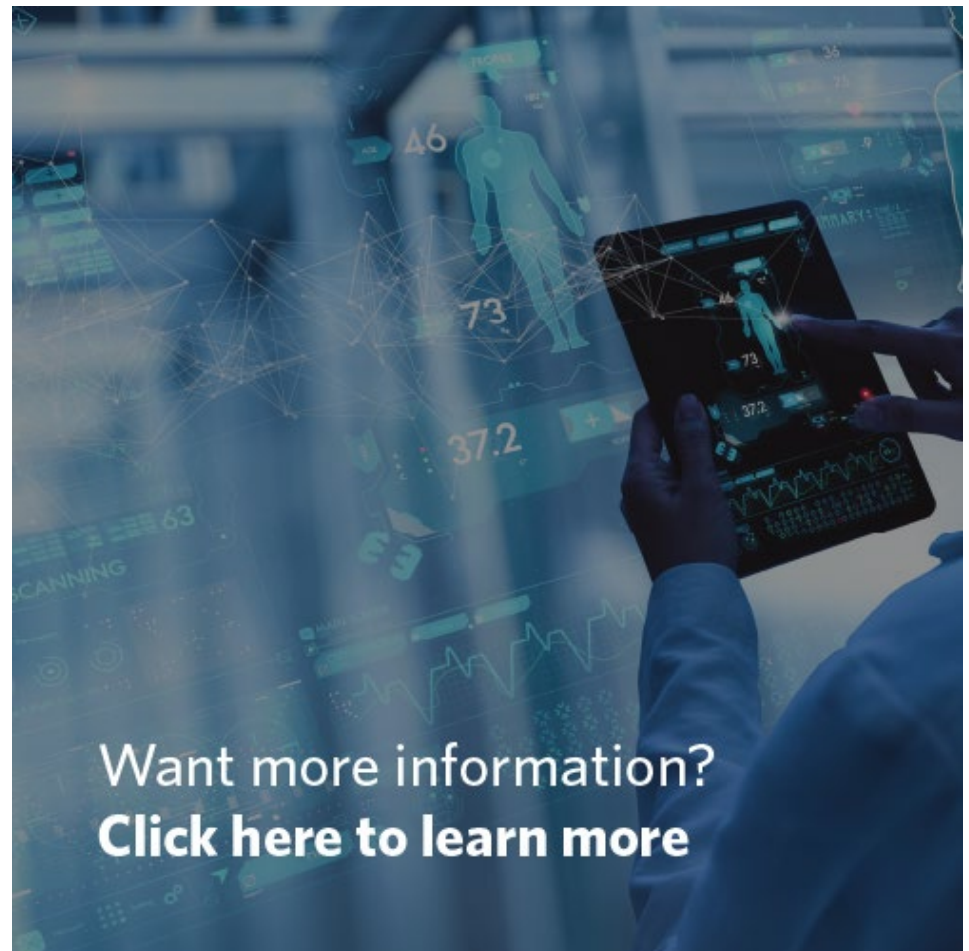


CONCLUSION

Predicting surgery duration is the first step to optimizing a surgeon's time. The development of these prediction models will help the hospital in better deciding the appropriate block of time required when making scheduling decisions. The next phase of work will involve using these models to develop an optimization strategy and solution, which will help the transformation team meet the goal of creating a better daily rhythm for surgeons.

Endnotes

1. <https://pubmed.ncbi.nlm.nih.gov/23351763/#:~:text=Results%3A%20In%20the%20100%20cases,underestimated%2032%25%20of%20the%20time.>



Want more information?
Click here to learn more

