mosaic
DATA SCIENCE
http://mosaicdatascience.com/
info@mosaicdatascience.com

Data Science Case Study

Demand Forecasting for Network Bandwidth

We built a custom demand forecasting model using ML techniques, helping a leading professional services firm predict office bandwidth requirements.

Industry

Management Consulting

Use Case

Predicting Office Bandwidth Needs

Techniques

Quantile and Linear Regression, Linear Mixed Effects Model, Simulation

Outcome

Interactive Web App powered by Machine Learning recommends bandwidth requirements days, weeks, and months out.

DEMAND FORECASTING BACKGROUND

One of the world's leading and top-rated management consulting firms was preparing to move their video conferencing and audio conferencing capabilities from an in-office hosted solution to Cisco WebEx. High-quality, seamless video and audio communication is critical to maintaining this firm's customer-facing reputation and to effective use of high-value company resources.

The firm's internal Information Technology (IT) team was tasked with facilitating a smooth transition, requiring that adequate network connectivity be established and available from each office to support the WebEx sessions which the firm conducts on top of other network traffic. The consulting firm has 85 offices in 48 countries with over 6,200 consultants and around 12,000 total staff. The firm believed that through analysis of their historical data along with predictive modeling they could build forecasts for how much bandwidth was needed at each office.

Mosaic was contracted as a strategic data science consultant to provide quantitative and predictive analyses of these systems to give the client confidence that after the transition the bandwidth would be sufficient to support the firm's daily operational needs. In addition, the management consulting firm also desired the ability to update results quickly and easily in order to monitor changes in usage, requirements, or other characteristics affecting the results.

Mosaic was able to draw upon its wealth of analytics consulting experience to provide an easily updateable web application that provides quick and informative analysis of the audio/video conferencing system allowing the customer to make data-driven decisions.

DEMAND FORECASTING DEVELOPMENT

Mosaic modeled the bandwidth requirements of the WebEx service, which is to be carried over dedicated network links. The modeling tasks included gaining access to, understanding, and organizing the data, predicting future bandwidth requirements based on current requirements, evaluating the accuracy of the predictive models, and documenting the tools created and the processes followed to create those tools. The main data sources for the model included employee count by hour by office, and detailed online meeting information.

The Mosaic team investigated a number of solutions including modeling the data with quantile regression and linear regression but taking the high estimate of the prediction interval. In the end, an ostensibly simple path was chosen: to remove up to the 50th percentile (depending on the amount of data available for each office) of the distribution of the number of concurrent meetings. This allowed the Mosaic team to approximate the 90th percentile as a high end estimation. This data was then modeled by a linear mixed effects (LME) model[1]. The LME model assumes a linear relationship between the number of people in the office and the number of concurrent meetings and allows this linear

relationship to vary by office and hour of the day. To evaluate the performance of this model the team performed the following steps:

1. For each office, all the available data was used to calculate the 90th percentile of the hourly number of concurrent meetings. This is the reference curve with which we compare the model output to evaluate performance.

2. For each office, all the available data was used to calculate the 90th percentile of the hourly number of employees for each hour of the day.

3. The model was evaluated using as input the employee counts obtained in step 2. The output of the model is the expected number of concurrent meetings

4. The curves obtained in steps 1 and 3 were compared.

In addition to creating the predictive model, the Mosaic team recognized that an interactive view of the historical data and predictive model would be very useful and informative to the customer for making data-driven decisions. To that end, the team used the R Shiny package to develop an interactive web application that allows viewers to see projected bandwidth needs at a chosen office under a given set of future operating conditions.

The Shiny app provides location-specific default operating conditions derived from the historical data when a user selects a location to view but also allows users to interactively adjust the parameters that define the operating conditions in order to see how differing assumptions impact predicted bandwidth usage. Example parameters include:

- Office Space Utilization – expected number of employees in the office on a given day
- Proportion Video – the proportion of video vs audio-only attendees of meetings
- Video Bandwidth – bandwidth required for video data
- Audio Bandwidth – bandwidth required for audio data
- Proportion Non-VOIP Audio – for example, calling in to a meeting with a cell phone

The interactivity enabled by the Shiny app allows stakeholders to adjust for expected vs actual changes in operating conditions (such as number of employees in an office) over time. The client plans to add data to the analysis on an ongoing basis so that they can continuously predict bandwidth requirements as the needs change.

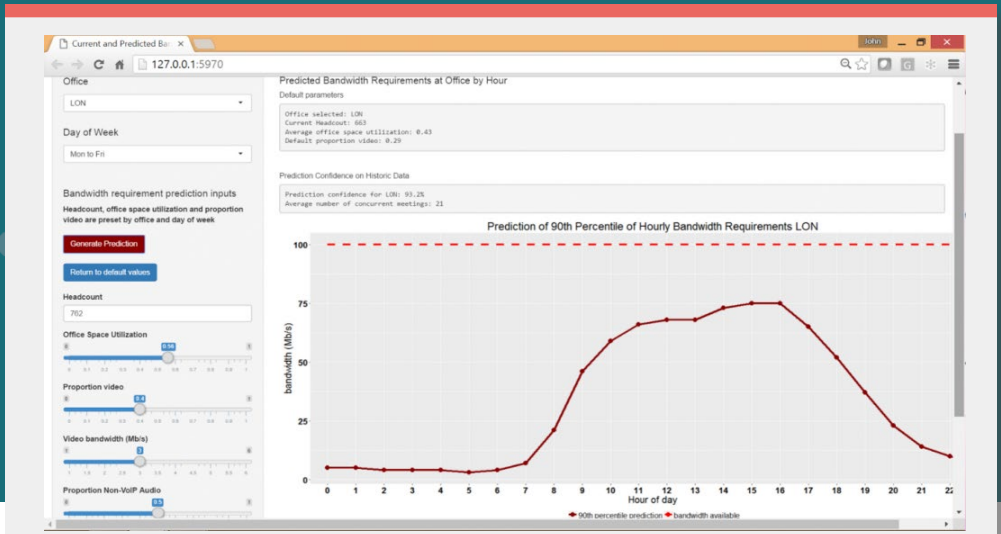
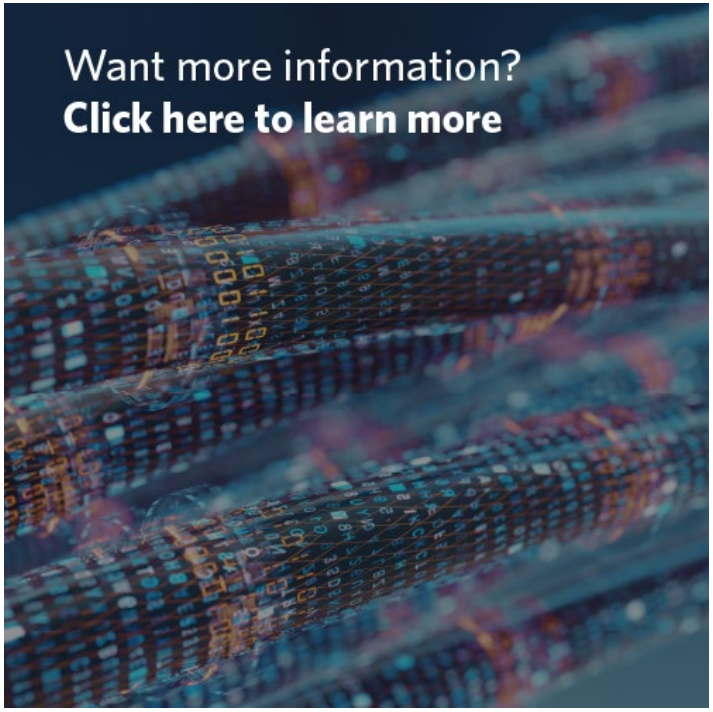


Figure 1: Predicted Bandwidth Requirements: high estimate (90th pct)

DEMAND FORECASTING RESULTS

Mosaic's detailed analyses and interactive web app gave the customer the capability to perform what-if analysis that were previously out of reach. The customer is now able to better understand how future changes will impact bandwidth needs at existing and new office locations. This allows for better and more efficient planning and use of resources, fewer dropped calls, and consistent high quality audio and video ensuring the firm's exceptional customer-facing reputation. Perhaps of even greater importance, Mosaic's analysis has changed how the customer thinks about these decisions. After seeing the results for the first time, one IT leader exclaimed, "We've never been able to make decisions in this way before!" How can analyses like this change the way you do business?



Endnotes

1. https://stats.idre.ucla.edu/other/mult-pkg/introduction-to-linear-mixed-models/