

Practical Data Science with R

Nina Zumel

John Mount

11

Producing effective presentations

This chapter covers

- Presenting your results to project sponsors
- Communicating with your model's end users
- Presenting your results to fellow data scientists

In the previous chapter, you saw how to effectively document your day-to-day project work and how to deploy your model into production. This included the additional documentation needed to support operations teams. In this chapter, we'll look at how to present the results of your project to other interested parties.

We'll continue with the example from last chapter: our company (let's call it WVCorp) makes and sells home electronic devices and associated software and apps. WVCorp wants to monitor topics on the company's product forums and discussion board to identify "about-to-buzz" issues: topics that are posed to generate a lot of interest and active discussion. This information can be used by product and marketing teams to proactively identify desired product features for future releases, and to quickly discover issues with existing product features. Once we've successfully built a model for identifying about-to-buzz topics on the forum, we'll want to

explain the work to the project sponsor, and also to the product managers, marketing managers, and support engineering managers who will be using the results of our model.

Table 11.1 summarizes the relevant entities in our scenario, including products that are sold by our company and by competitors.

Table 11.1 Entities in the buzz model scenario

| Entity | Description |
|--------------|---|
| WVCorp | The company you work for |
| eRead | WVCorp's e-book reader |
| TimeWrangler | WVCorp's time-management app |
| BookBits | A competitor's e-book reader |
| GCal | A third-party cloud-based calendar service that TimeWrangler can integrate with |

A disclaimer about the data and the example project

The dataset that we used for the buzz model was collected from Tom's Hardware (tomshardware.com), an actual forum for discussing electronics and electronic devices. Tom's Hardware is not associated with any specific product vendor, and the dataset doesn't specify the topics that were recorded. The example scenario we're using in this chapter was chosen to present a situation that would produce data similar to the data in the Tom's Hardware dataset. All product names and forum topics in our example are fictitious.

Let's start with the presentation for the project sponsors.¹

11.1 Presenting your results to the project sponsor

As we mentioned in chapter 1, the project sponsor is the person who wants the data science result—generally for the business need that it will fill. Though project sponsors may have technical or quantitative backgrounds and may enjoy hearing about technical details and nuances, their primary interest is business-oriented, so you should discuss your results in terms of the business problem, with a minimum of technical detail.

You should also remember that the sponsor will often be interested in “selling” your work to others in the organization, to drum up support and additional resources to keep the project going. Your presentation will be part of what the sponsor will share

with these other people, who may not be as familiar with the context of the project as you and your sponsor are.

To cover these considerations, we recommend a structure similar to the following:

- 1 Summarize the motivation behind the project, and its goals.
- 2 State the project's results.
- 3 Back up the results with details, as needed.
- 4 Discuss recommendations, outstanding issues, and possible future work.

Some people also recommend an “Executive Summary” slide: a one-slide synopsis of steps 1 and 2.

How you treat each step—how long, how much detail—depends on your audience and your situation. In general, we recommend keeping the presentation short. In this section, we'll offer some example slides in the context of our buzz model example.

Let's go through each step in detail.

We'll concentrate on content, not visuals

In our discussion, we'll concentrate on the content of the presentations, rather than the visual format of the slides. In an actual presentation, you'd likely prefer more visuals and less text than the slides that we provide here. If you're looking for guidance on presentation visuals, a good book is *The Craft of Scientific Presentations* by Michael Alley (Springer, 2003).

If you peruse that text, you'll notice that our bullet-laden example presentation violates all his suggestions. Think of our skeleton presentations as outlines that you'd flesh out into a more compelling visual format.

It's worth pointing out that a visually oriented, low-text format like Alley recommends is meant to be *presented*, not read. It's common for presentation decks to be passed around in lieu of reports or memos. If you're distributing your presentation to people who won't see you deliver it, make sure to include comprehensive speaker's notes. Otherwise, it may be more appropriate to go with a bullet-laden, text-heavy presentation format.

11.1.1 Summarizing the project's goals

This section of the presentation is intended to provide context for the rest of the talk, especially if it will be distributed to others in the company who weren't as closely involved as your project sponsor was. Let's put together the goal slides for the WVCorp buzz model example.

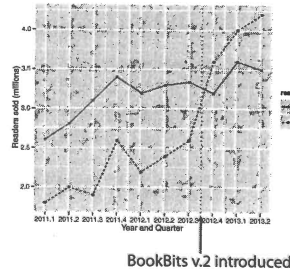
In figure 11.1, we provide background for the motivation behind the project by showing the business need and how the project will address that need. In our example, eRead is WVCorp's e-book reader, which led the market until our competitor released a new version of their e-book reader, BookBits. The new version of BookBits has a shared-bookshelves feature that eRead doesn't provide—though many eRead users expressed the desire for such functionality on the forums. Unfortunately, forum

¹ We provide the PDF versions (with notes) of our example presentations at <https://github.com/WinVector/zmPDSwR/tree/master/Buzz> as ProjectSponsorPresentation.pdf, UserPresentation.pdf, and PeerPresentation.pdf.

If Only We'd Known...

eRead vs. BookBits

- eRead: Best selling indie ebook reader - until BookBits v.2
- Hot new BookBits feature: shared bookshelves
- Our one-at-a-time book lending didn't compete
- Estimated \$25M lost revenue on product sales



Show the business need that motivated this project: WVCorp lost revenue and market share to a competitor.

Solid curve: WVCorp's product (eRead)
Dashed curve: Competitor's product (BookBits)

Could We Have Caught This?

- eRead Forum discussions:
 - Sharing a booklist with a friend, to grab from as they pleased
 - Sharing a book with a group of friends (first-come-first-serve)
- Whenever these questions arose, the discussion was lively
 - Suggestions, work-arounds, kludges, "me too"s
 - A shared bookshelf (like BookBits) would have met these recurring needs
- There was Buzz around this issue! But we ignored it. Or didn't find it.
 - Labor intensive to continually keep up with forum activity

WVCorp has the information to help address this need, but not enough resources (labor) to use the information effectively.

In a real presentation, use a screenshot of a relevant forum discussion.

Figure 11.1 Motivation for project

traffic is so high that product managers have a hard time keeping up, and somehow missed detecting this expression of users' needs. Hence, WVCorp lost market share by not anticipating the demand for the shared-bookshelf feature.

In figure 11.2, we state the project's goal, in the context of the motivation that we set up in figure 11.1: we want to detect topics on the forum that are about to buzz so that product managers can find emerging issues early.

Once you've established the project's context, you should move directly to the project's results. Your presentation isn't a thriller movie—don't keep your audience in suspense!

11.1.2 Stating the project's results

This section of the presentation briefly describes what you did, and what the results were, in the context of the business need. Figure 11.3 describes the buzz model pilot study, and what we found.

Goal: Catch it Early

- Predict which topics on our product forums will have persistent buzz
- Features customers want
- Existing features users have trouble with
- Persistent buzz, not ephemeral or trendy issues
- Persistence = real, ongoing customer need

State the project goal in terms of the business motivation.

In a real presentation, use a screenshot of a relevant forum discussion.

Figure 11.2 Stating the project goal

Pilot Study

- Collected three weeks of data from forum
- Trained model on Week 1 to identify which topics will buzz in Weeks 2/3
 - Buzz = Sustained increase of 500+ active discussions in topic/day, relative to Week 1, Day 1
- Compared predicted results to topics that actually buzzed
- Feedback from team of five product managers—how useful were the results?

Briefly describe how the project was run.

Results

- Reduced manual scan of forums by over a factor of 4
- Scan 184 topics—not 791!
- PMs: 75% of identified topics produced "valuable insight"
- Found 84% of about-to-buzz topics
- Low (20%) false positive rate

| | Predicted No Buzz | Predicted Buzz | |
|---------|-------------------|----------------|-----|
| No Buzz | 579 | 35 | 614 |
| Buzz | 28 | 149 | 177 |
| Total | 607 | 184 | 791 |

State the results up front.

State the results in terms of how they affect the end users (product managers).

The model reduces the end users' workload by zeroing in on what they need to look at.

Representative end users thought the model's output was useful.

Figure 11.3 Describing the project and its results

Keep the discussion of the results concrete and nontechnical. Your audience isn't interested in the details of your model *per se*, but rather in why your model will help solve the problem that you stated in the motivation section of the talk. Don't talk about your model's performance in terms of precision and recall or other technical metrics, but rather in terms of how it reduced the workload for the model's end users, how useful they found the results to be, and what the model missed. In projects where the model is more closely tied to monetary outcomes, like loan default prediction, try to estimate how much money your model could potentially generate, whether as earnings or savings, for the company.

11.1.3 Filling in the details

Once your audience knows what you've done, why, and how well you've succeeded (from a business point of view), you can fill in details to help them understand more. As before, try to keep the discussion relatively nontechnical and grounded in the business process. A description of where the model fits in the business process or workflow and some examples of interesting findings would go well in this section, as shown in figure 11.4.

The “How it Works” slide in figure 11.4 shows where the buzz model fits into a product manager's workflow. We emphasize that (so far) we've built the model using metrics that were already implemented into the system (thus minimizing the number of new processes to be introduced into the workflow). We also introduce the ways in which the output from our model can potentially be used: to generate leads for potential new features, and to alert product support groups to impending problems.

The bottom slide of figure 11.4 presents an interesting finding from the project (in a real presentation, you'd want to show more than one). In this example, TimeWrangler is WVCorp's time-management product, and GCal is a third-party cloud-based calendar service that TimeWrangler can talk to. In this slide, we show how the model was able to identify an integration issue between TimeWrangler and GCal sooner than the TimeWrangler team would have otherwise (from the customer support logs). Examples like this make the value of the model concrete.

We've also included one slide in this presentation to discuss the modeling algorithm (shown in figure 11.5). Whether you use this slide depends on the audience—some of your listeners may have a technical background and will be interested in hearing about your choice of modeling methods. Other audiences may not care. In any case, keep it brief, and focus on a high-level description of the technique and why you felt it was a good choice. If anyone in the audience wants more detail, they can ask—and if you anticipate such people in your audience, you can have additional slides to cover likely questions. Otherwise, be prepared to cover this point quickly, or to skip it altogether.

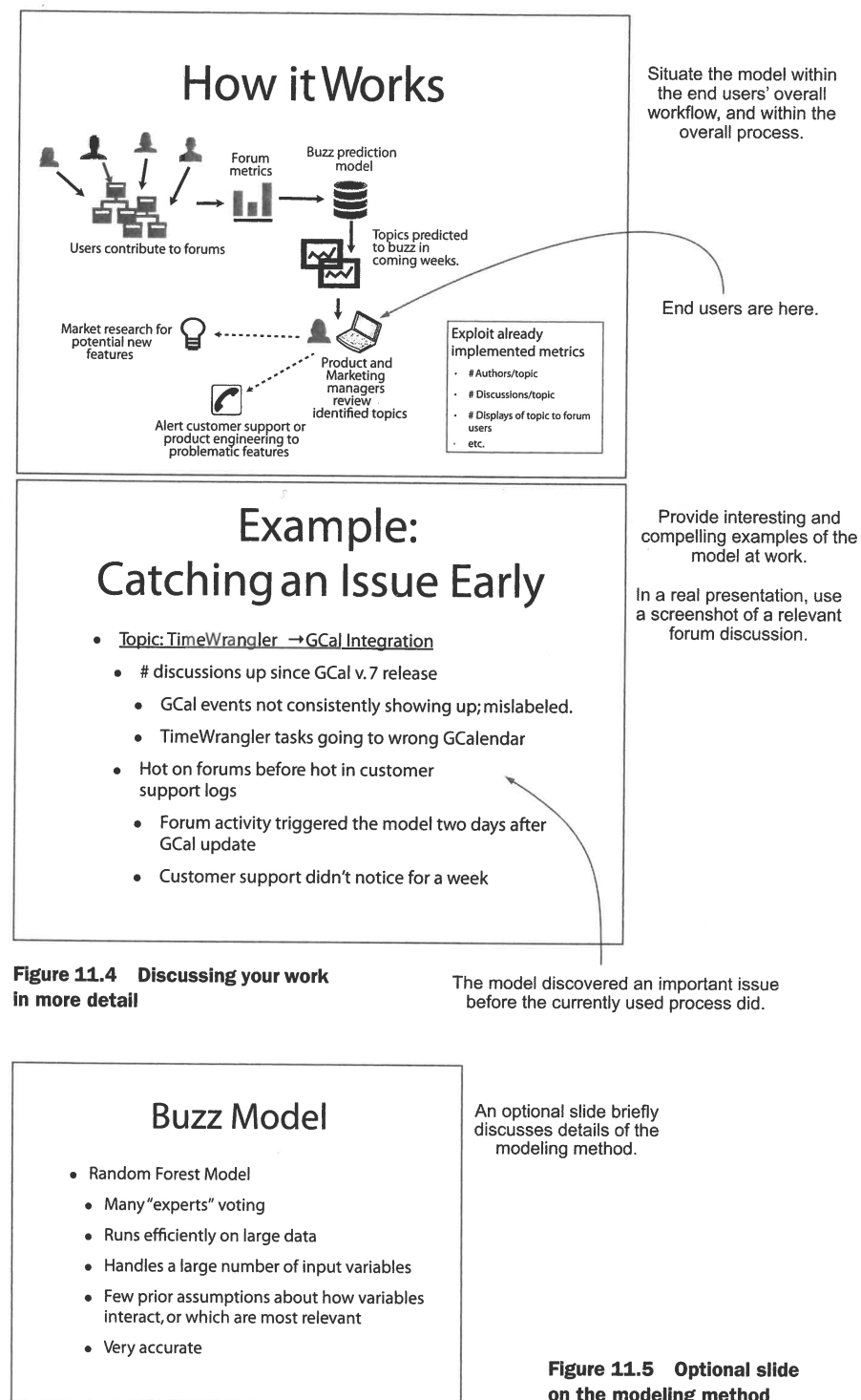


Figure 11.4 Discussing your work in more detail

The model discovered an important issue before the currently used process did.

Figure 11.5 Optional slide on the modeling method

There are other details that you might want to discuss in this section. For example, if the product managers who participated in your pilot study gave you interesting quotes or feedback—how much easier their job is when they use the model, findings that they thought were especially valuable, ideas they had about how the model could be improved—you can mention that feedback here. This is your chance to get others in the company interested in your work on this project and to drum up continuing support for follow-up efforts.

11.1.4 *Making recommendations and discussing future work*

No project ever produces a perfect outcome, and you should be up-front (but optimistic) about the limitations of your results. In the buzz model example, we end the presentation by listing some improvements and follow-ups that we'd like to make. This is shown in figure 11.6. As a data scientist, you're of course interested in improving the model's performance, but to the audience, improving the model is less important than improving the process (and better meeting the business need). Frame the discussion from that perspective.

The project sponsor presentation focuses on the big picture and how your results help to better address a business need. A presentation for end users will cover much of the same ground, but now you frame the discussion in terms of the end users' workflow and concerns. We'll look at an end user presentation for the buzz model in the next section.

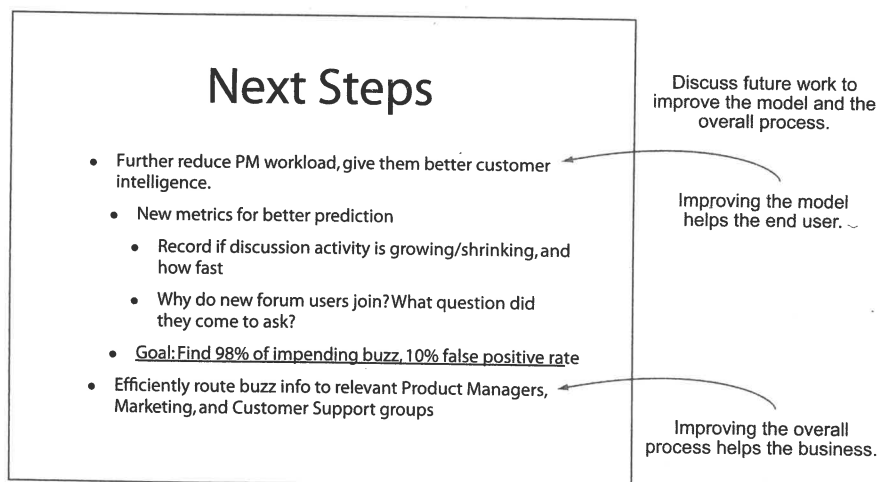


Figure 11.6 Discussing future work

11.1.5 *Project sponsor presentation takeaways*

Here's what you should remember about the project sponsor presentation:

- Keep it short.
- Keep it focused on the business issues, not the technical ones.
- Your project sponsor might use your presentation to help sell the project or its results to the rest of the organization. Keep that in mind when presenting background and motivation.
- Introduce your results early in the presentation, rather than building up to them.

11.2 *Presenting your model to end users*

No matter how well your model performs, it's important that the people who will actually be using it have confidence in its output and are willing to adopt it. Otherwise, the model won't be used, and your efforts will have been wasted. Hopefully, you had end users involved in the project—in our buzz model example, we had five product managers helping with the pilot study. End users can help you sell the benefits of the model to their peers.

In this section, we'll give an example of how you might present the results of your project to the end users. Depending on the situation, you may not always be giving an explicit presentation: you may be providing a user's manual or other documentation. However the information about your model is passed to the users, we believe that it's important to let them know how the model is intended to make their workflow easier, not more complicated. For the purposes of this chapter, we'll use a presentation format.

For an end user presentation, we recommend a structure similar to the following:

- 1 Summarize the motivation behind the project, and its goals.
- 2 Show how the model fits into the users' workflow (and how it improves that workflow).
- 3 Show how to use the model.

Let's explore each of these points in turn, starting with project goals.

11.2.1 Summarizing the project's goals

With the model's end users, it's less important to discuss business motivations and more important to focus on how the model affects them. In our example, product managers are already monitoring the forums to get a sense of customers' needs and issues. The goal of our project is to help them focus their attention on the "good stuff"—buzz. The example slide in figure 11.7 goes directly to this point. The users already know that they want to find buzz; our model will help them search more effectively.

Our Goal: Catch User Needs Early

- Predict which topics on our product forums will have persistent buzz
- Features customers want
- Existing features users have trouble with
- Persistent buzz, not ephemeral or trendy issues
- Persistence = real, ongoing customer need

Provide motivation for the work from the end user's perspective: help them find useful buzz faster.

In a real presentation, you might use a screenshot of a relevant forum discussion.

Figure 11.7
Motivation for project

11.2.2 Showing how the model fits the users' workflow

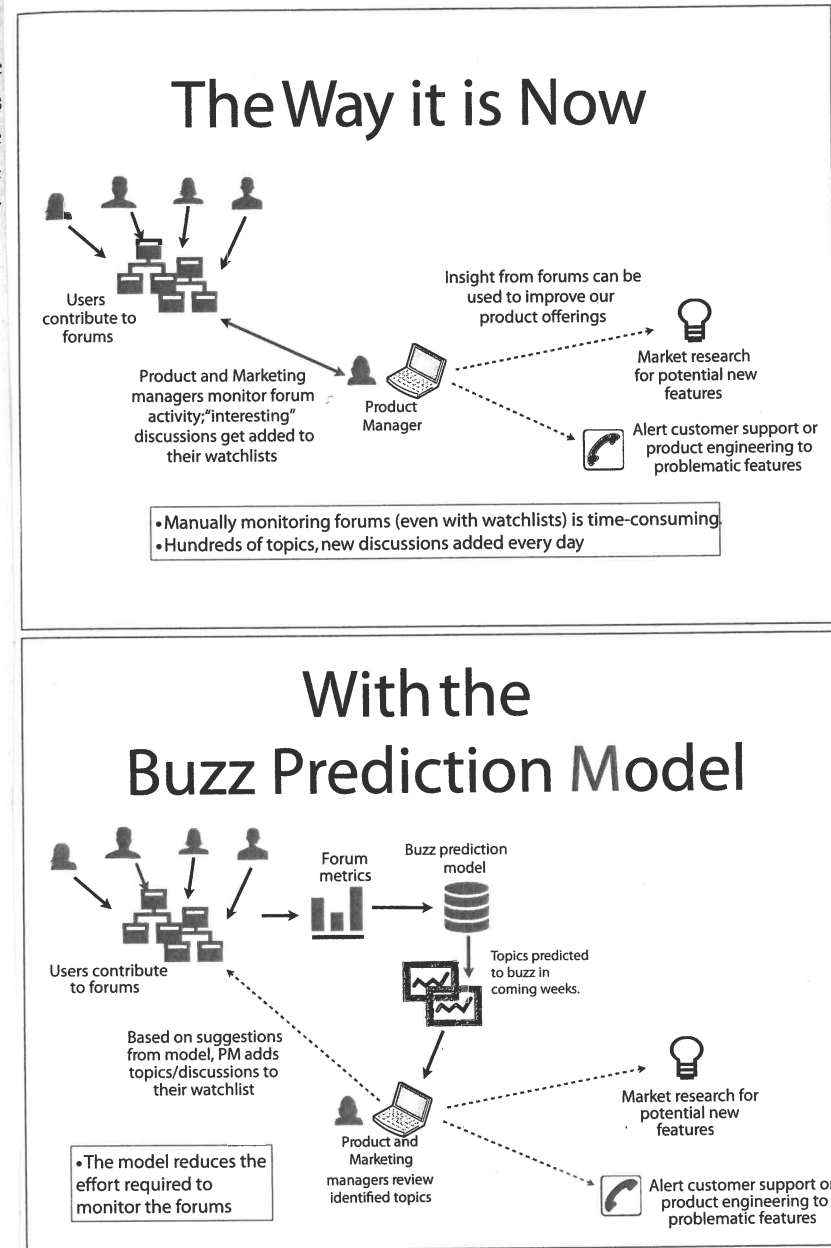
In this section of the presentation, you explain how the model helps the users do their job. A good way to do this is to give before-and-after scenarios of a typical user workflow, as we show in figure 11.8.

Presumably, the before process and its minuses are already obvious to the users. The after slide emphasizes how the model will do some preliminary filtering of forum topics for them. The output of the model helps the users manage their already existing watchlists, and of course the users can still go directly to the forums as well.

The next slide (figure 11.9, top) uses the pilot study results to show that the model can reduce the effort it takes to monitor the forums, and does in fact provide useful information. We elaborate on this with a compelling example in the bottom slide of figure 11.9 (the TimeWrangler example that we also used in the project sponsor presentation).

You may also want to fill in more details about how the model operates. For example, users may want to know what the inputs to the model are (figure 11.10), so that they can compare those inputs with what they themselves consider when looking for interesting information on the forums manually.

Once you've shown how the model fits into the users' workflow, you can explain how the users will use it.

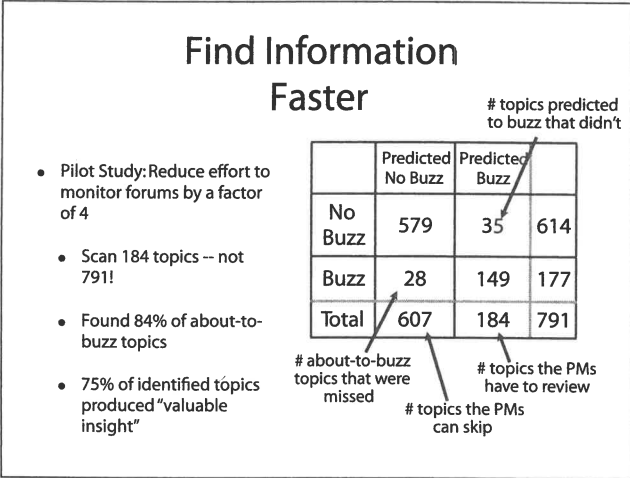


Compare the end users' day-to-day work process before and after the introduction of your model.

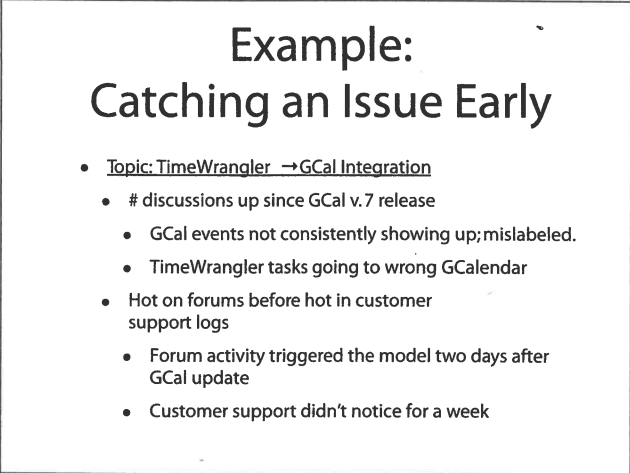
Before: time-consuming.

After: more focused, less time-consuming.

Figure 11.8 User workflow before and after the model



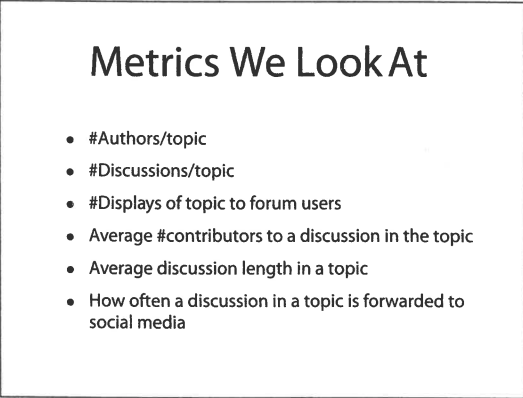
State the results from the end user's perspective: manual effort is reduced, and the model's suggestions are correct and valuable.



Show interesting and compelling examples of the model at work.

In a real presentation, you might use a screenshot of a relevant forum discussion.

Figure 11.9 Present the model's benefits from the users' perspective.

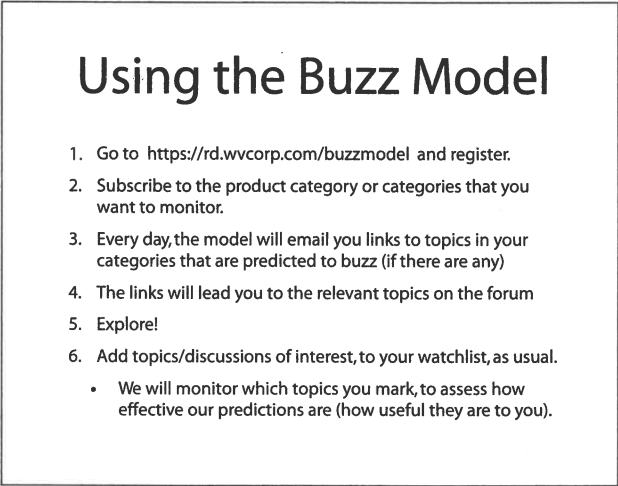


The end users will likely be interested in the inputs to the model (to compare with their own mental processes when they look for buzz manually).

Figure 11.10 Provide technical details that are relevant to the users.

11.2.3 Showing how to use the model

This section is likely the bulk of the presentation, where you'll teach the users how to use the model. The slide in figure 11.11 describes how a product manager will interact with the buzz model. In this example scenario, we're assuming that there's an existing mechanism for product managers to add topics and discussions from the forums to a watchlist, as well as a way for product managers to monitor that watchlist. The model will separately send the users notifications about impending buzz on topics they're interested in.



Show the users how to interact with the model.

In a real presentation, each point would be expanded with step-by-step detailed instructions and appropriate screenshots.

Figure 11.11 Describe how the users will interact with the model.

In a real presentation, you'd then expand each point to walk the users through how they use the model: screenshots of the GUIs that they use to interact with the model, and screenshots of model output. We give one example slide in figure 11.12: a screenshot of a notification email, annotated to explain the view to the user.

By the end of this section, the user should understand how to use the buzz model and what to do with the buzz model's output.

Finally, we've included a slide that asks the users for feedback on the model, once they've been using it in earnest. This is shown in figure 11.13. Feedback from the users can help you (and other teams that help to support the model once it's operational) to improve the experience for the users, making it more likely that the model will be accepted and widely adopted.

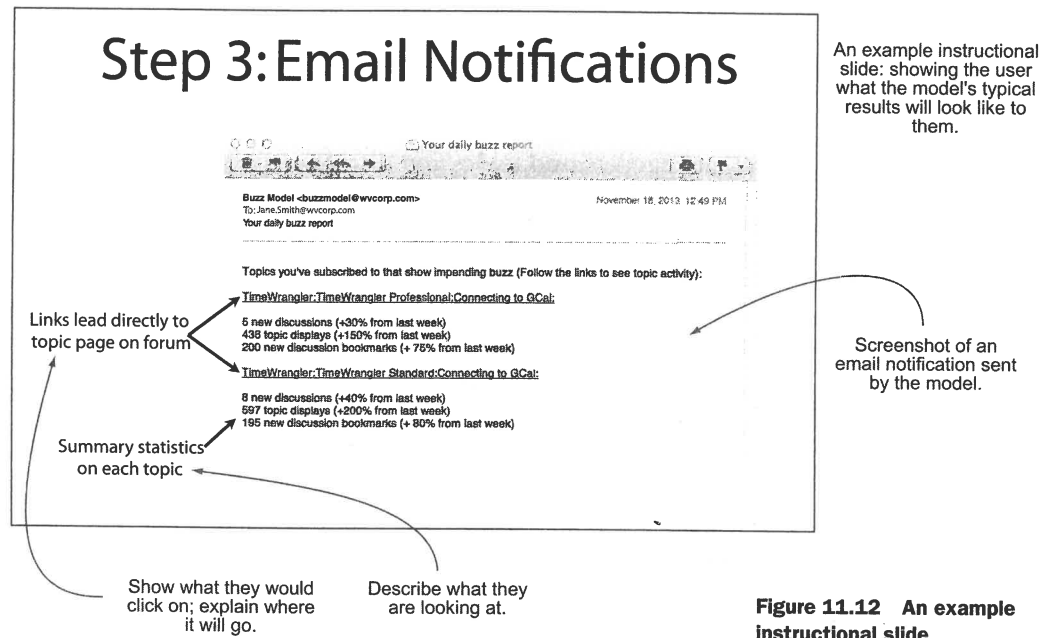


Figure 11.12 An example instructional slide

11.3 Presenting your work to other data scientists

Presenting to other data scientists gives them a chance to evaluate your work and give you a chance to benefit from their insight. They may see something in the problem that you missed, and can suggest good variations to your approach or alternative approaches that you didn't think of.

Other data scientists will primarily be interested in the modeling approach that you used, any variations on the standard techniques that you tried, and interesting findings related to the modeling process. A presentation to your peers generally has the following structure:

- 1 Introduce the problem.
- 2 Discuss related work.
- 3 Discuss your approach.
- 4 Give results and findings.
- 5 Discuss future work.

Let's go through these steps in detail.

11.3.1 Introducing the problem

Your peers will generally be most interested in the prediction task (if that's what it is that you're trying to solve, and don't need as much background about motivation from the project sponsors or the end users. In figure 11.14, we start off by introducing the concept of buzz and why it's important, then go straight into the prediction task.

This approach is best when you're presenting to other data scientists within your own organization, since you all share the context of the organization's needs. When you're presenting to peer groups outside your organization, you may want to lead with the business problem (for example, the first two slides of the project sponsor presentation, figures 11.1 and 11.2) to provide them with some context.

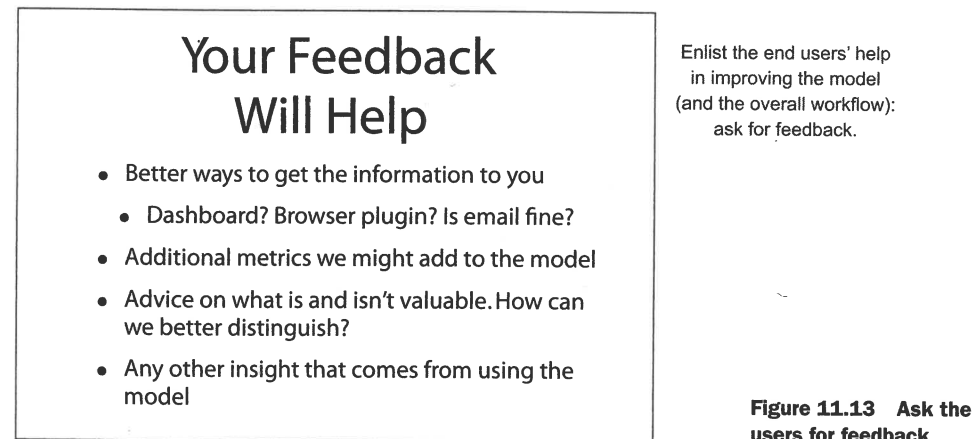


Figure 11.13 Ask the users for feedback.

In addition to presenting your model to the project sponsors and to end users, you may be presenting your work to other data scientists in your organization, or outside of it. We'll cover peer presentations in the next section.

11.2.4 End user presentation takeaways

Here's what you should remember about the end user presentation:

- Your primary goal is to convince the users that they want to use your model.
- Focus on how the model affects (improves) the end users' day-to-day processes.
- Describe how to use the model and how to interpret or use the model's outputs.

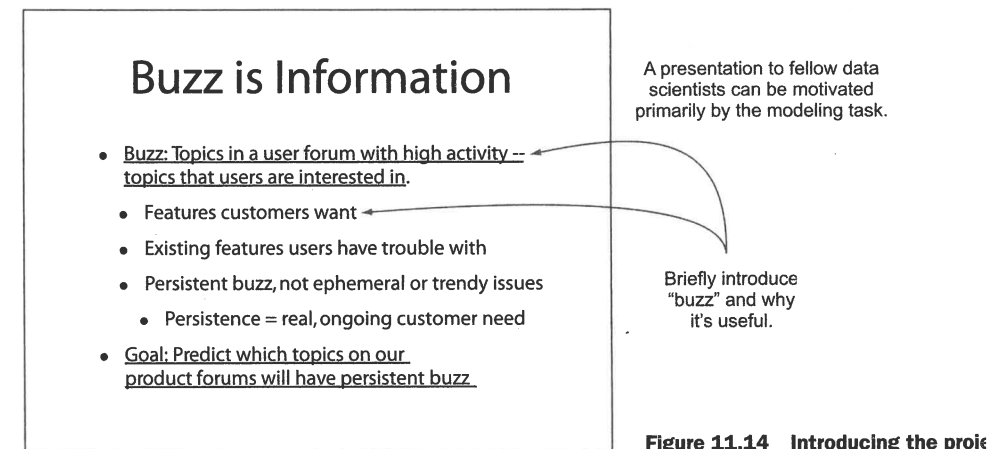


Figure 11.14 Introducing the project