Effective BI Visualization for AI Prediction

Part 3, Text Analysis

December 2023

By BI & AI Committee

Just as Sally Sue was getting a solid grasp on visualizing the predictions and the forecasts from the organization's structured data, lo and behold they asked her to start using artificial intelligence for its unstructured data. Phoooey! A lesser person might panic, but she's the legendary Sally Sue so she dug right into text topic analysis, sentiment analysis, and emotion analysis.

Topic Analysis

Text topic analysis is a text mining technique where large amounts of text are analyzed and summarized into one or more topics/ themes. You may have an enormous amount of string data within a database, such as large, unstructured paragraphs of information. These may be reviews, emails, technical support tickets, or anything that comes from free-form text input. There is

no specific target to predict like in a traditional AI model. Instead, you want to understand what the main topics within the text are. Think of these topics like general themes. If you could break down a large amount of text, how would you describe it in a few words? For example, suppose you have a dataset that consists of:

- A product

- A review title

- A review date

- A rating from 1-5

- Free-form review text; each review is generically known as a document

Product	Review Title	Review Date	Stars	Review Text
Xbox	Love it!	01/03/2006	5	I have waited all year to buy this. The graphics are so good!
PS3	Best Purchase!	04/03/2007	5	The cell processor is second to none. Highly recommended!
Wii	Great Console!	07/04/2008	5	I love the motion control sticks, they're so great!

With enough text, you can use an algorithm to extract topics and terms from the reviews and assign one or more topics to them. Suppose Sally Sue is tasked with analyzing these reviews and trying to understand both the general topics and sentiment within them. She must run a text topic model on the above data. Her

algorithm will associate one or more generalized topics with each review. A topic can be thought of as basic, few-word theme of the text.

| Product	Review Title	Review Date	Stars	Review Text	_Topics
Xbox	Love it!	01/03/2006	5	I have waited all year to buy this. The graphics are so good!	Graphics, Good, Love
PS3	Best Purchase!	04/03/2007	5	The cell processor is second to none. Highly recommended!	Blu-Ray, Movies, Realistic
Wii	Great Console!	07/04/2008	5	I love the motion control sticks, they're so great!	Motion, Controllers, Unique

While the algorithms to arrive at these results are complex and require a great deal of iteration, the overall method to visualize the results are simple. There are three primary visualizations Sally uses that are helpful for constructing a data story with text topic models:

1. A bar chart of topics and document count

2. A word cloud of terms

3. A table holding topics and other relevant data

Each of these is described in detail below.

1. Bar Chart

Knowing what topics are in her data and how often they occur is a key piece of information for Sally's stakeholders. A simple bar chart of topics by document count can give a wealth of information without being overwhelming. In this example, each document is a single review. Sally creates a bar chart that indicates the total number of reviews for each topic.
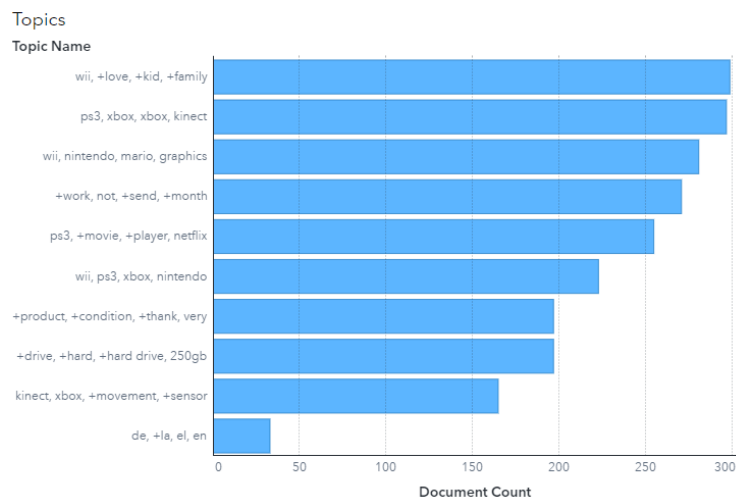
She identifies 10 topics that are important and uses the bar chart to show how many reviews contain those topics. When identifying

the number of topics to display, aim for the minimum number of topics needed to tell a useful story and create actionable insight. A high number of topics may become overwhelming for the end user viewing the report. For example, consider the additional complexity added to the analysis below by increasing the amount to 25 topics:
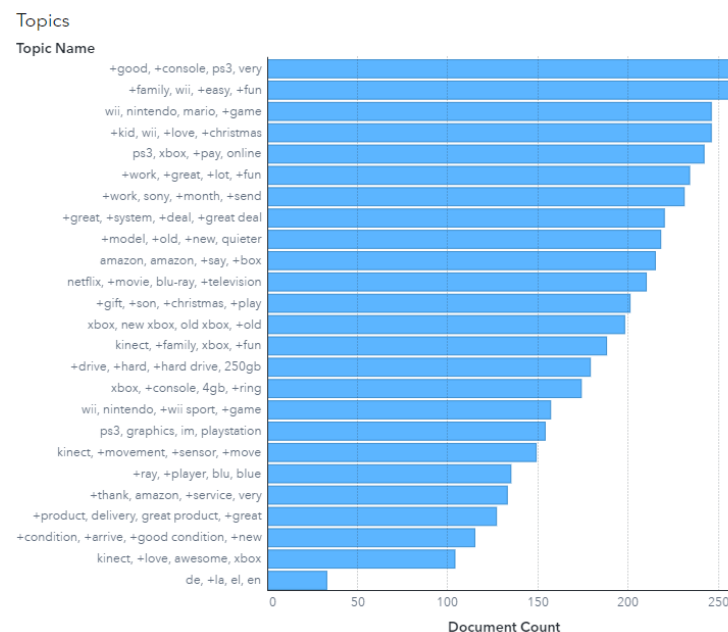
While this is only a general guideline, you should take your unique business situation into account to identify the total number of relevant topics needed to solve your specific problem or tell your business story. This may very well be hundreds of topics. In general, try to minimize the number of topics displayed so that it is easy for a business analyst or key stakeholder to absorb them.

2. Word Count

Text topic analysis typically includes term analysis, where important terms are extracted from the text. Since these terms are one word only, Sally found that a word cloud is a great way to help

users identify the most important terms within the text. The size of the word within the cloud is determined by the number of times that the term occurs within the collection of reviews. Important terms are big and readily stand out.

She also adds a filter interaction between the topic bar chart and terms word cloud to make it even more powerful, enabling users to see which terms are within each topic. For example, selecting a topic in the bar chart filters the word cloud to display only terms from reviews which contain that topic.

3. Topics Table

When dealing with raw text, tables of raw data are crucial. Sally's end-users want to be able to review the raw text themselves, narrow down topics, or even do a text search for certain terms. There are two types of topics tables that Sally uses to better enable her end-users:

1. An interactive table that is filterable to a specific topic and term

2. A searchable detailed table that can be exported for further analysis

The interactive table is on the same page as the bar chart and word cloud, and it interacts with both the bar chart and word cloud. For example, selecting a topic on the bar chart filters the

list table to reviews with that topic, and selecting a term from the word cloud subsequently filters the list table to only include reviews that contain that term. This table should include, at minimum, the raw document text. Sally further populates it with relevant data that is helpful for a business analyst. In this example, she includes the review title, rating, review date, and raw review text.

Business analysts or even data scientists may want to do additional analysis from this data. As a result, Sally creates a separate page holding a table of terms and documents that can be exported and used for additional analysis. Each review may have multiple topics associated with it. The most logical way of handling this is to use a binary indicator for each topic: 1 if a topic is within a review, and 0 if a topic is not. While slightly less user-friendly than "Yes" or "No," this is incredibly easy for one of Sally's business analysts or fellow data scientists to export and work with externally.

Sally found that using cell highlighting is a simple but effective way to make this table more user friendly within her business intelligence application while enabling business analysts or data scientists to export the data and perform additional analyses. Note the search bar above the table. This allows a user to find all text in the table that contains a word or phrase. When including the ability to search text, make sure it is case-insensitive and consider including typeahead/autocomplete capabilities.

Putting it all Together

With a bar chart, word cloud, and document table, Sally's users can get a great deal of information. To make it more useful, she adds interactions between all of these graphics so that users can easily filter values. For example, selecting a topic in the bar chart can filter both the word cloud and the document table, and selecting a term in the word cloud can further filter the document table.
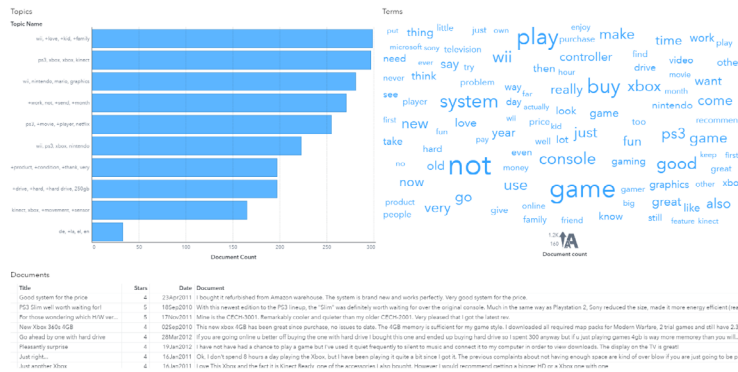
Text topic analysis is a very complex field, but its visualizations do not need to be. A few simple graphs and a little bit of interactivity can enable business analyst to dive into their freeform text data and summarize it in a way that tells a compelling data story.

Sentiment Analysis

We began this series of posts by sharing how Sally Sue's eyes were opening to the fact that binary classifications weren't always so binary. Yet we left you hanging by never focusing on how to visualize those binary (0/1, Yes/No, True/False) type of predictions. But don't worry, we will resolve that for you in this section as we tackle how Sally Sue was learning to handle another type of AI type that can be performed on things like survey comments and social media posts called Sentiment Analysis.

It involves passing a passage like "Nursing staff was incredibly rude despite their excellent clinical care" to a Machine Learning Model trained for Natural Language Understanding (NLU) of some flavor, and having it let you know if the comment/response was Positive or Negative. So, a binary classification of sorts, but with cooler sounding words than Yes/No or True/False.

To give you a real feel for the complexity of these positives or negatives from sentiment analysis let's try our hand at analyzing the sentiment of the example statement ourselves. "Incredibly rude" sure is negative. But "excellent clinical care" sure seems positive. So, if you were acting as an AI model, how would you respond? That's not all that easy. Is it?

As you know the AI models don't just return the binary "positive" or "negative" they also return a score. So, while you are scratching your head feel free to respond with a 3-digit score of how confident you are in your positive or negative response. Not all that easy. Is it?

Guess what? One sentiment analysis model (using a 0 to 1 scale) categorized it as ".967 positive, .025 negative and .008 neutral." Almost 100% positive. While another categorized it as -.6 negative on a -1 to 1 scale, so very negative. Clearly models, like human beings, have a bias coded into them that more highly values certain words. Undoubtedly different models for predictions based on structured data would return different scores, but more than likely two models wouldn't completely diverge on whether a customer was going to churn by such a wide range.

Not sure what you want to do with that information, but for the rest of this section we will just ignore it. Instead, for the remainder of this section just assume that we scored five statements with a single AI sentiment analysis model. The text of the statements themselves are meaningless compared to the importance of visualizing the results. Our selected sentiment analysis model returned the following values:

Direct Return from AI

Statement ID	Returned phrase
One	Negative
Two	Negative
Three	Negative
Four	Positive
Five	Positive

If you think that looks very similar to the non-binary, binary classification predictions then you are correct. More importantly, now you understand how vital it is for Sally Sue to visualize the whole story, and not just the words themselves.

She knew she would need a way to incorporate the values. Let's look at the actual returns with values as well. Because as with our categorical predictions the end user at minimum needs to see if the Negatives and Positives are strong or not.

Direct Return from AI

Statement ID	Returned phrase	Returned Value
One	Negative	-.65
Two	Negative	-.72
Three	Negative	-.85
Four	Positive	.51
Five	Positive	.95

Please note: The model returns values in a range 0-1. Can you imagine seeing a Negative .65 for one statement, while seeing a Positive .65 for another statement? So, we have flipped the sign for the values when the returned phrase is negative to help the end user distinguish and avoid that confusion.

Oh great, now she was back in the same situation. The words are clearly depicted, the numbers are clearly depicted, and she flipped the signs to avoid confusion. But as with predictions for structured data, Sally Sue felt like she needed to take another step to help the end user avoid doing the mental math and categorize the values to help them quickly comprehend the values.

Categorized

| Statement ID	Adjusted Phrase
One	Probably Negative
Two	More than likely Negative
Three	Very likely Negative
Four	Seriously it's a coin flip
Five	Definitely Positive

Within seconds we're sure you found fault with the word choices and came up with some of your own. That's excellent. It means you understand your end users and that you understand the value in the way the Sentiment Analysis, which is essentially a binary classifier for unstructured text, is displayed visually for those end users. The goal is to change the words positive or negative into phrases that combine the word and the score so that your users are confident doing something about them.

Going further, is there a more visual method that could be utilized?

We wouldn't have asked if we didn't think so. While some sentiment models return both positive/negative/neutral values, many only return a single word value along with the score. But the

other value is then implied. After all, we saw that with the very first example. Right?

Sally Sue sure thought so and tried her hand at using a Butterfly Chart to depict both the Negative and Positive values.
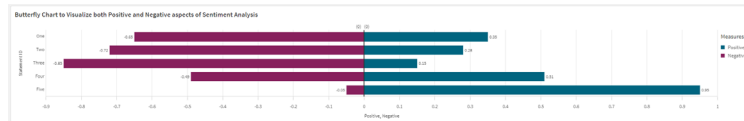

Sally Sue liked the fact that regardless of what word choices "strongly positive" or "coin flip" might be in a table, the strong ones are visually distinct and stand out quickly. Which is exactly what Business Intelligence Visualizations should do. But she wasn't happy with the display of the "Seriously, it's a coin flip" value. She wanted a visual that allowed her users to quickly ascertain the strong positive/negative values as well as recognize quickly what the coin flips were.

She lost a lot of sleep then began questioning what .51 "Positive" actually means, and how she could help her users remove all doubt.

"One way to remove the doubt from users might be to realize that the .51 positive value simply implies that it is only leaning .01 to the positive side of the coin?

So, she used adjusted values by removing .50 from each. The .51 was adjusted to .01, the .95 was adjusted to .45, etc. By doing so, she only needed to visually display the value for the word since the oppositive values were then adjusted back to a 0. Now the strongly scored values were obvious, while the coin flips were also obvious.
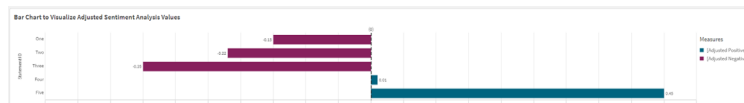
Another way to "remove the doubt" is to subtract out the doubt. Meaning, that there is .49 negative doubt that the sentiment was positive. So, in the chart below Sally Sue made those adjustments. She also decided to use some fancy schmancy gradient coloring so that the strength of the values was not only reflected in the length
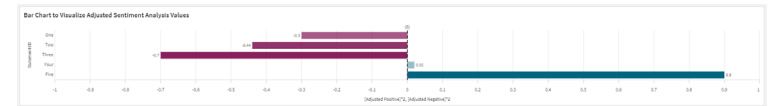


of the lines, but in the coloring of them as well.

Do you agree that adjusting the values to "remove the doubt," either by subtracting .50 or by subtracting the detracting values makes sense?

Do you think the gradient coloring adds value to indicate the confidence?

Do you think it might be better to have color bands that match the text classifications?

Do you think it might be better to use the color bands when displaying the text values in the tables so that "Seriously, it's a coin flip" shows up in a weak color, while "Definitely" shows up in a strong color?

Unlike you, Sally Sue was fortunate to never have more than 5 different surveys/comments to show at any 1 time. Just kidding. Naturally, just like the rest of us, she had to figure out ways to visualize much larger quantities of values. Like any other kind of number, this adjusted sentiment score could be visualized as an average and a median.
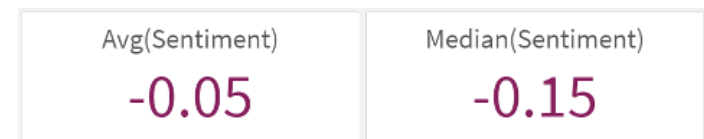

But like with any measure, those kinds of aggregates provide the

"answer" but not necessarily the whole story. She tried a box plot.
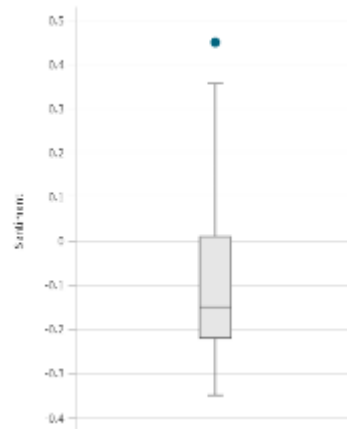
It showed what it was supposed to show visually, but she wasn't in love with the fact her end users might not be comfortable with such a data "sciencey" type of visualization. Sure the .45 positive "outlier" stood out as statistically significant to her, but would others understand that?

Then she tried her hand with a Distribution Plot which shows each sentiment score so you can see how they are "distributed."
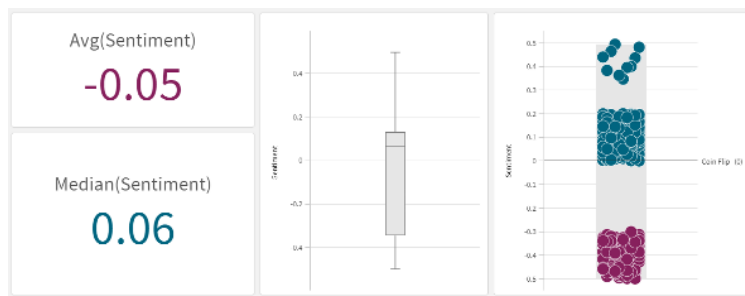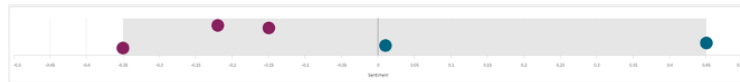
Do the KPI's, the Box Plot and the Distribution Plot all show the exact same set of values? Absolutely. But notice how the story unfolds moving left to right. The KPI's tell us that more values are on the positive side than the negative side because the Median is positive, even though the average is negative. For those that understand Box Plots that is also clear. But when the Distribution Plot is added, the story really unfolds: "There are more positive values, but the bulk of them are only up to .2 positive, while only a handful are above .3. Yet when the sentiment is negative it is highly negative." In other words, customers/responders either kind of like us, or really don't like us.

This would be a great time to refresh your memory that Business Intelligence isn't intended to just turn numbers into visuals. It is meant to help drive decisions and move people to action. We can see an executive's response now:

"Ok Sally Sue, I see that a lot of our customers have a pretty negative sentiment. But what are we supposed to do about it?"

In her gut, Sally Sue knew they had no idea what to do to improve the sentiment. All the visualizations simply showed them a sentiment score. At the end of the day, no matter how you visualize the words positive or negative, sentiment analysis will leave your end users feeling a bit in want.

The good news is that while Sentiment Analysis is the most common way to classify unstructured textual statements, it's far from the only artificial intelligence model that can be used or visualized. So, Sally Sue took it as an action item to dig into some others, to see if they might offer information about the statements that did provide a deeper level of insight.

Emotional Text Classification is one such other model. Very similar to the text classification we already looked at, and as the name suggests, these models attempt to report information about the suggested emotions of the words, beyond just telling you that the overall sentiment is positive or negative. For our statement

"nursing staff was incredibly rude despite their excellent clinical care" one such emotional text classification model yielded the following results:
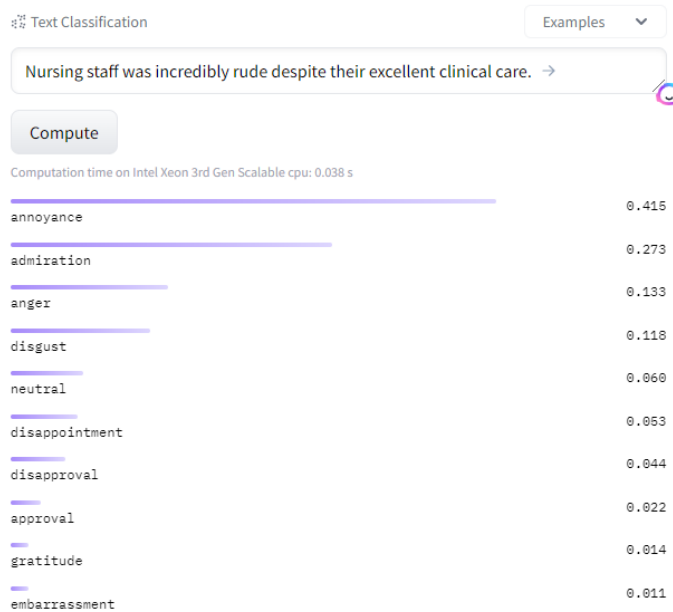
So much to unpack from these results. First, let's focus on the sniff test. Do you agree with the results? Do you think the patient who wrote that comment on their survey probably felt annoyed at the nursing staff? Odd we know, but do you believe that while being annoyed the patient also felt admiration for the nursing staff?

We think that the dichotomy of emotions is exactly what the patient was feeling, and is so much more powerful than just reporting the word positive or negative.

Now the question begging to be asked is, "How we would visualize the words returned from emotional text classification?" Oh yeah, just like we did for the base text classification models, in a word cloud.

Since we are only visualizing the results for a single statement the size of the word is based on the strength of the emotion. However, it is just as easy to size the words based on how many times they were referenced. Imagine Sally Sue's audience when they see the following word cloud and understand that it is based on 383 user surveys.

Regardless, how we visualized the word negative, or what strength it had, can you imagine the difference in their hearts when they see this word cloud? While they still may not know what specific

actions to take, they will be driven to want to act. To us, that is the magical power of artificial intelligence combined with business intelligence. Instead of just visualizing a number, we are now visualizing emotions and helping the reader feel empathy and want to act, not just read.

To drive deeper, we need to stop being so generic in our usage. The statement we looked at as an example is clearly healthcare-related. So, let's acknowledge that Sally Sue is working for a health system. The survey comments she is trying to analyze are from something known as a HCAHPS Survey, which stands for Hospital Consumer Assessment of Healthcare Providers and Systems.

What if we could literally ask for a breakdown of the statement along the 5 measurements that HCAHPS concerns itself? That would be crazy awesome right?

Generative AI

You know that hype we said we would avoid about Generative AI, well let's play along, just this once to see what it might offer us. Instead of just seeing sentiment, or just seeing emotion, what if we could ask a Generative AI engine to specifically score our statement based on the HCAHPS measures. That's just a pipe dream. Unless it isn't:

Can you give me a numerical sentiment analysis value for the following for each of the HCAPS categories "Nursing staff was incredibly rude despite their excellent clinical care. "

Numerical sentiment analysis values are subjective and can vary depending on the specific sentiment analysis model used. However, I can provide you with a general sentiment analysis interpretation of the text you provided. In this case, I will use a scale of -1 to +1, where -1 represents extremely negative sentiment, 0 represents neutral sentiment, and +1 represents extremely positive sentiment.

1. Communication with Nurses: -0.8
2. Responsiveness of Hospital Staff: -0.7
3. Care Transition: -0.5
4. Pain Management: 0.0 (not mentioned in the text)
5. Overall Rating and Recommendation: -0.7

These numerical values indicate negative sentiment in the HCAPS categories based on the text provided. The author expresses dissatisfaction with the nursing staff, describing them as incredibly rude despite acknowledging their excellent clinical care. The negative sentiment suggests a perception of poor communication, lack of responsiveness, and potentially negative interactions with the nursing staff.

Unlike the sentiment analysis that can lean more towards "incredibly rude" or "excellent clinical care" this analysis isn't swayed by the "excellent clinical care" because that isn't an HCAHPS category. This analysis provides us with exactly what Sally Sue's user groups would need to see. The Sentiment Analysis shown visually would tell them they need to improve. The Emotional Analysis would stir their heart to want to improve. Assuming we aggregated up the scores and didn't fly off the handle over this one result, the aggregated analysis would suggest that they should focus on improving communications and responsiveness.

If you have ever been in a relationship, undoubtedly you understand the key to a good relationship is communication. But how do you improve communication? While we are embracing the hype of Generative AI, instead of dancing around the subject let's just shoot for the stars and ask for ideas on how to improve.

So, you know, that is not edited in any way shape or form. That was the legitimate response from a Generative AI tool. Clearly some of the steps are already occurring in health systems worldwide. But some of the suggestions are profound.

Conclusion

This would be crazy ... what if we constructed a Business Intelligence dashboard that:

1. Visualized the sentiment analysis so that leadership knew thedirection customers/patients/staff were leaning.

2. Visualized the emotional classification results to help them empathize with the customers/patients/staff to know why they were leaning that way.

3. Visualized and aggregated suggested improvements to help them quickly take action.

The purpose of modern Business Intelligence is to do far more than just show numbers, it is to drive people to act. The natural progression for Sally Sue was to depict ways for her business users to play "what-if" games with the data. "What-if I changed this, would it help us move in a more positive direction? What-if we did that, would it increase our customer retention?"

Authors

Cupid Chan

Cupid Chan, an industry veteran, holds the Senior Fellow position at CDHAI in Johns Hopkins and the University of Maryland. His career began in developing a top-tier BI platform and contributing to Linux Foundation ODPi. As Chair of LF AI & Data's BI & AI Committee, he drives the fusion of AI and BI, forging Cognitive Intelligence (CI), melding AI's speed with human-driven insights.

Deepak Karuppiah

Deepak Karuppiah, a Senior Architect in MicroStrategy's Augmented Analytics Group, merges AI and BI to craft technology stacks for AI-powered dashboards. Formerly, he specialized in secure data connectors for MicroStrategy. Proficient in machine vision and learning, he explores novel AI-BI intersections, evident through his role in the BI & AI committee. Beyond work, he volunteers technical skills to a local non-profit, showcasing dedication beyond his professional sphere.

Dalton Ruer

Dalton is a Senior Solution Architect at Qlik. He is a Data Scientist Storyteller, Analytics Evangelist and is an impassioned student of Generative AI. He is a seasoned author, speaker, blogger and YouTube video creator who is best known for dynamically sharing inconvenient truths and observations in a humorous manner. The passion which Dalton shares through all mediums, moves and motivates others to action.

Sachin Sinha

Sachin, Microsoft's Director of Technology Strategy, has honed his data engineering expertise post-graduation from the University of Maryland. Specializing in information management, he crafted systems aiding data-driven decision-making. His support extended to healthcare startups, fostering data-centric business models, and empowered public sector entities toward their missions via data-enabled decisions. Residing in Fairfax, VA, with his family, he passionately backs the Terps and Ravens while excelling in his professional pursuits.

Stu Sztukowski

Stu, a senior product manager at SAS, focuses on user-friendly AI through visual analytics. With a BS in Statistics from North Carolina State University and an MS in Advanced Analytics, he transitioned from data science, excelling in forecasting and BI. He champions accessible high-performance AI, aiming for universal comprehension. A mentor and versatile leader, he simplifies complex analytics for data scientists and business analysts, driven by a passion for public speaking and democratizing analytics.

LF AI & DATA

About the LF AI & Data Foundation

LF AI & Data is an umbrella foundation of the Linux Foundation that supports open source innovation in artificial intelligence (AI) and data. LF AI & Data was created to support open source AI and data, and to create a sustainable open source AI and data ecosystem that makes it easy to create AI and data products and services using open source technologies. We foster collaboration under a neutral environment with an open governance in support of the harmonization and acceleration of open source technical projects.Explore our current portfolio of projects and contact us to discuss hosting your open source AI or data project under our Foundation.

December 2023