Data 650 – Fall 2019

Assignment 4 – Ravens User Engagement Chatbot

Kenneth Lulie

Professor Purushothman

University of Maryland University College

November 30, 2019

**Creation of Chatbot “The Raving Raven” to Increase Fan Engagement**

**Introduction**

The purpose of the chatbot created for this assignment, the “Raving Raven”, is to increase user engagement among fans of the Baltimore Ravens football team. The requirements are that users are able to answer and receive questions regarding Ravens trivia as well as provide information for users on where to watch the next Ravens Game. An end-user would be Raven’s fans or potential Raven’s fans. The end user will interact with the bot either through integration with the Raven’s website, or through an app. The expected benefits would be to increase user engagement with the fan base. The Ravens as a sports team relies on fans being engaged and interested in the team and due to the network effect an increase in engagement in one fan can often lead to increases in engagement with other people as well. Increased user engagement leads to increased advertising metrics and increased numbers of Raven’s merchandise being sold both of which are essential to long-term success of the Ravens as a commercial enterprise.

FlowChart and Dialog Scenarios

The chatbot is designed to always start at the “Welcome” node where the app welcomes the user, describes the app, and asks for the users name (See Figure 1). The welcome node then always sends the user to the “Central” node, where the app welcomes the user by name and then lists the services available in the app. The user can then go to the “Players”, “Games”, “Stats”, or “Where to Watch” nodes.

If the “Players” intent is detected, the chatbot will send the user to the Players node. The node welcomes the user, and offers to provide information on star Raven’s football players. Currently, three players are coded in as entities under the player entities. Each of these players has a separate child node that the chatbot will send the user to if the player’s entity is detected and the child node will tell the user information about that Raven’s player. The players are the Ravens Quarterback Lamar Jackson, their star Running Back, Mark Ingram, and their star wide receiver, Willie Sneed. This node increases user engagement because often fans will pick favorite players, and the narrative of each player is a powerful story often used to enhance drama and involvement with the team.

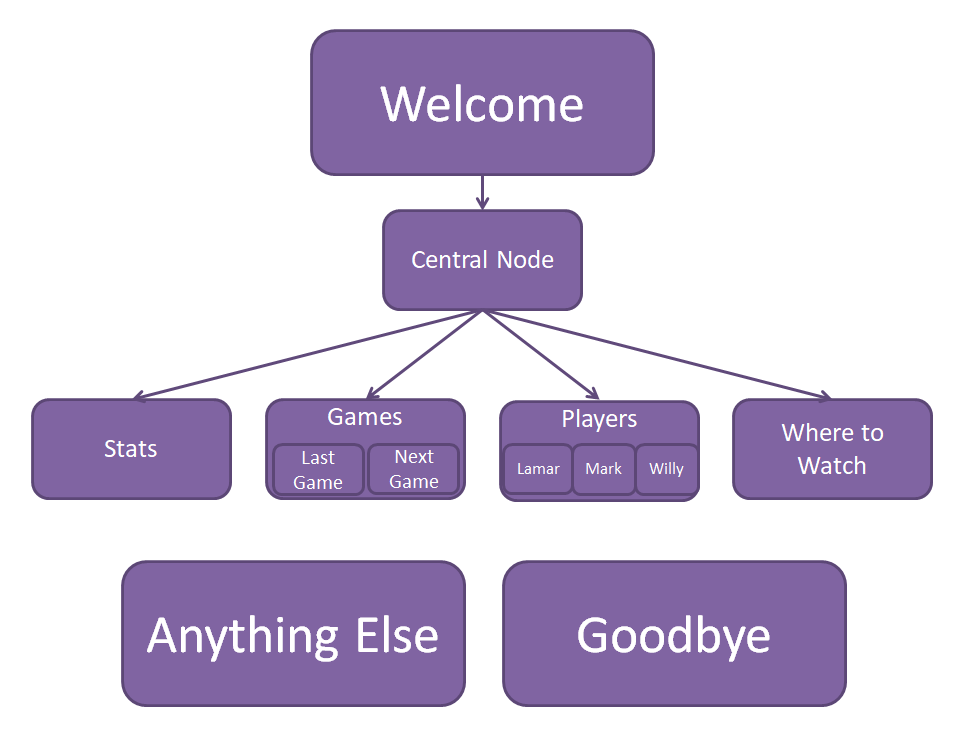
If the chatbot detects the “Games” intent, it will send the user to the “Games” node. The node will welcome the user, and list as options information about the last game the Ravens played, or information about the next game the Ravens will play. Each of these options has a child node the user will be sent to if they pick the respective option, either provided the results of the last game and who it was played against, or when the next game is played, and against whom among other information. This node increases user engagement by providing the user with information needed to stay up to date with the team and discuss with other fans.

If the chatbot detects the “Stats” intent the chatbot will send the user to the “Stats” node where the node will welcome the user and provide the user with statistics about how the ravens as a whole have performed so far in the season. Statistics are often a way that fans will understand performance year to year, player to player, or team to team and are a prime way to help users stay engaged and involved with what the team is doing.

The last major node is the “Where\_to\_Watch” node, which is where the user will be sent to if the “where\_to\_watch” intent is detected. It is important as it informs the user as to where they are able to watch the ravens on TV, listen on radio, or watch through a streaming service. The viewership rights for NFL games are very valuable and often changing hands year to year, and the involvement of streaming services often adds complexity to it as well. By providing the user with all the information they need on where to watch it increases the odds the user will actually view the game increasing engagement.

The “AnythingElse” node is where the user is sent if no known intents are detected. The node informs the user the chatbot was unable to understand the request, and lists the options available. There is also a “Goodbye” node for when the “Goodbye” intent is detected which thanks the user for using the app and exits the chatbot, or would in an actual implementation.

Figure 1. The Raving Raven Chatbot Dialog Flowchart



The chatbot was developed according to the specifications above. Intents were created for each of the nodes as described above, specifically Players, Games, Stats, Where to Watch, and Goodbye with each having at least 5 synonyms, and entities were created with the parent category of Games having last game and next game and the parent category of Players having an entity for each of the three players. At least five synonyms were provided for each of these entities.

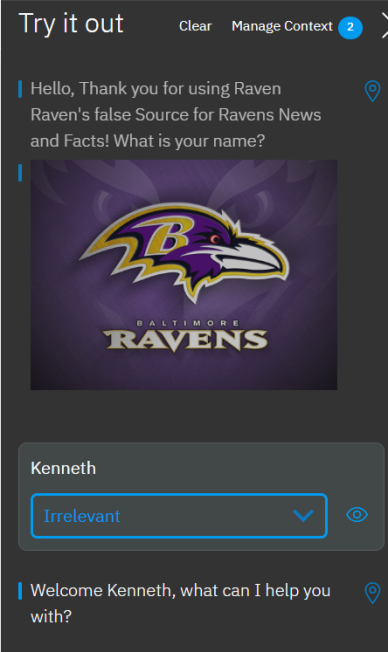
The welcome node was designed to prompt for the name of the user which is used in the central node and the anything else node was designed to be a helpful way for the user to get back on track to the content the chatbot has to offer.

**Chatbot Dialog Testing**

**Welcome & Central Nodes**

As expected, the Welcome and Central nodes worked with the user name being stored and then used (See **Figure 2**). This is not an area that should be having a large amounts of errors due to the simplicity of the flow so it was encouraging to see the expected behavior.

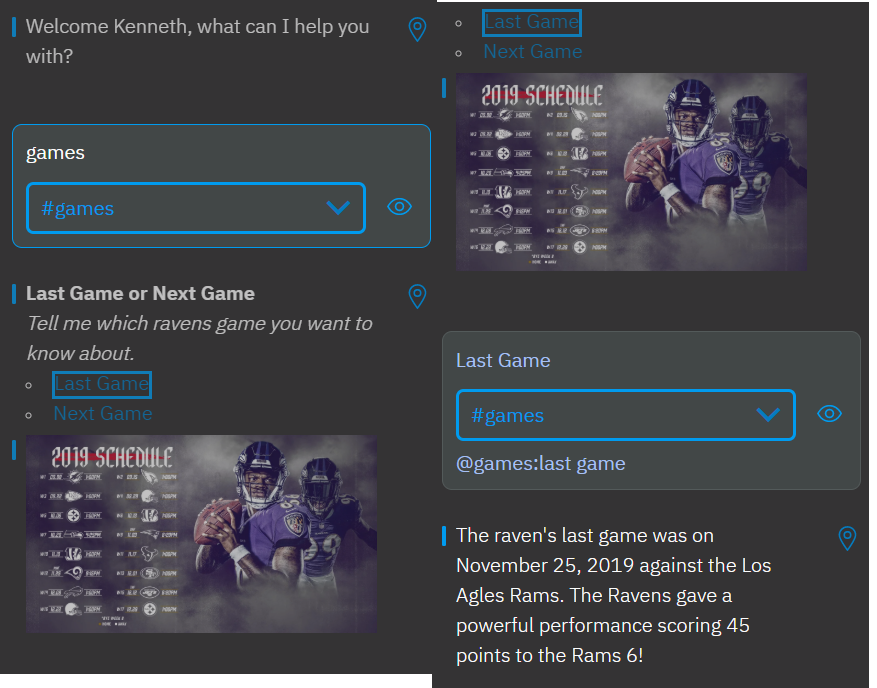
Figure 2. Welcome and Central Nodes Testing



**Games Node Testing**

The games node was tested by first seeing if the games intent was properly detected which it was, and then if the user was sent to the games node which they were. Then the ‘Last Game’ option was selected, and as intended the user was given information regarding the last game (See **Figure 3**). The next game option was also tested and worked as expected.

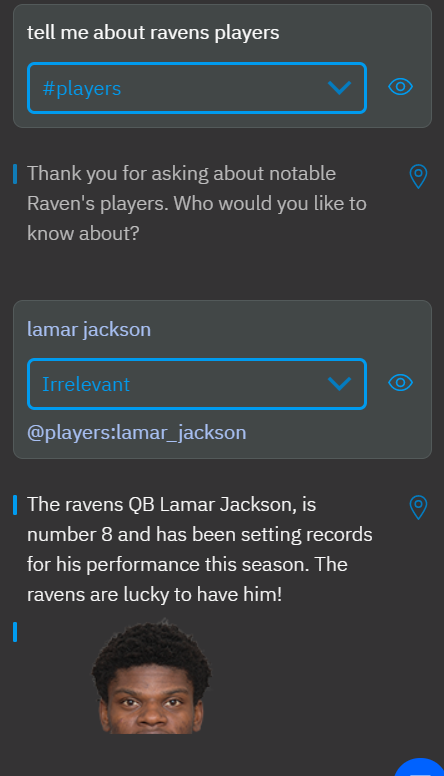
Figure 3. Games and Last Games Node Testing



**Players Node Testing**

The players node was tested to see if the players intent would be detected, which it was, and if the user would be sent to the players node which they were. Then one of the child nodes was tested by seeing if the chatbot would detect the players:lamar Jackson entity which was properly detected and did send the user to the lamar Jackson child node where information and a picture of the player was sent (**Figure 4**). The other two player entities were also tested and both worked as expected also providing information and a picture of the player.

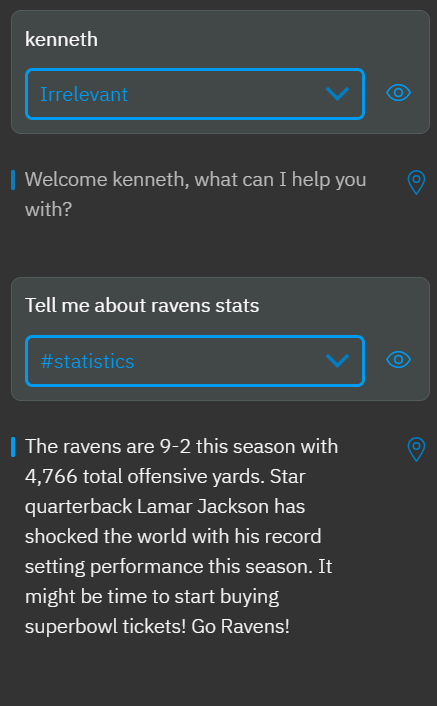
Figure 4. Players node and Lamar Child Node



**Stats Node Testing**

To test the stats node, the chatbot was tested to see if it would detect the stats intent, and then if the user would be sent to the stats node which they were **(See Figure 5)**. The chatbot then provided stats on the Ravens as well as a message interpreting the stats and engaging with the user by sending a positive message reading the Raven’s chances of making it to the superbowl.

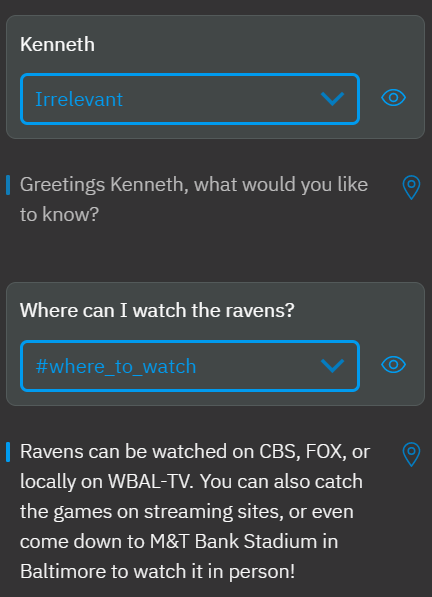
Figure 5. Stats Node Testing

****

**Where to Watch Node Testing**

The where to watch node was testing by seeing if the chatbot detected the where to watch intent which it did, and then if the user was sent to the where to watch node which they were. The node then provided information to the user regarding where they can view or listen to the next Ravens game as intended **(See Figure 6)**.

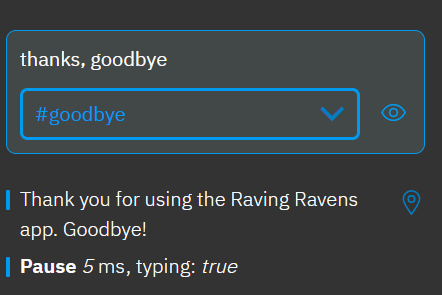
Figure 6. Where to Watch Node Testing

****

**Goodbye Node Testing**

The goodbye node was tested by seeing if the goodbye intent was detected which it was, and if the user was sent to the goodbye node which they were. The goodbye message was displayed and a pause is used to simulate the ending of the interaction.

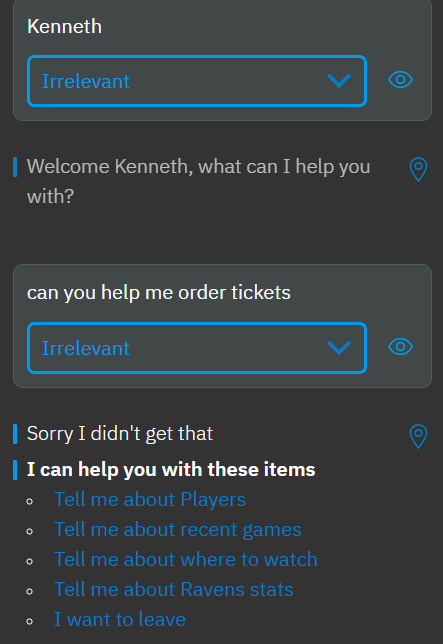
Figure 7. Goodbye Node Testing

****

**Anything Else Node Testing**

The anything else node was tested by asking the virtual assistant for help in buying tickets. As this does match any entered entity or intent the user was directed to the “Anything Else” node as intended. The anything else node then gave the user a list of options of what the app could help with as intended.

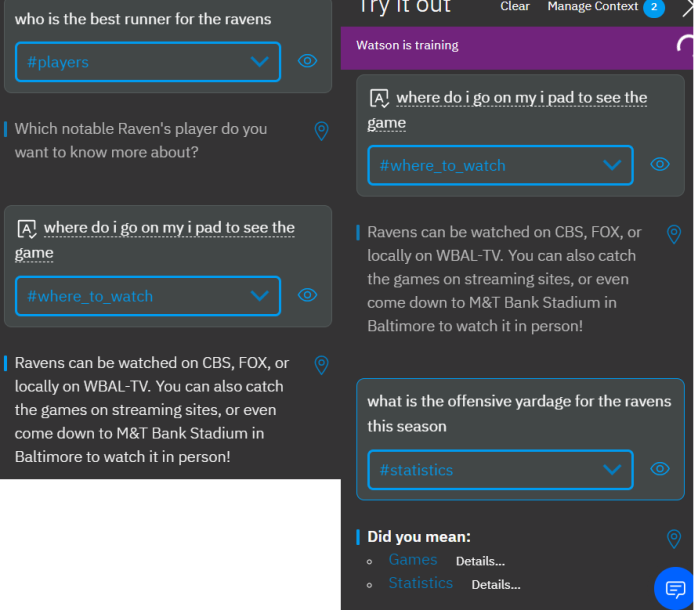
Figure 8. Anything Else Node Testing



**Testing and Training Virtual Assistant**

The virtual assistant was tested by running different dialog scenarios with wording similar but different to the listed synonyms to see how well the intent could be detected. In examples where the intent or entity was not properly detected, the virtual assistant was corrected using the try it out panel to correct the detection and the virtual assistant was retrained. Please see examples of this below **(Figure 9)**. Overall Virtual Assistant did a better job than expected of understanding the user input even when similar wording hadn’t been entered or trained. For example, it understood that a request about percentage of running yards being related to the statistics intent even though no wording was shared with the entered synonyms for that intent.

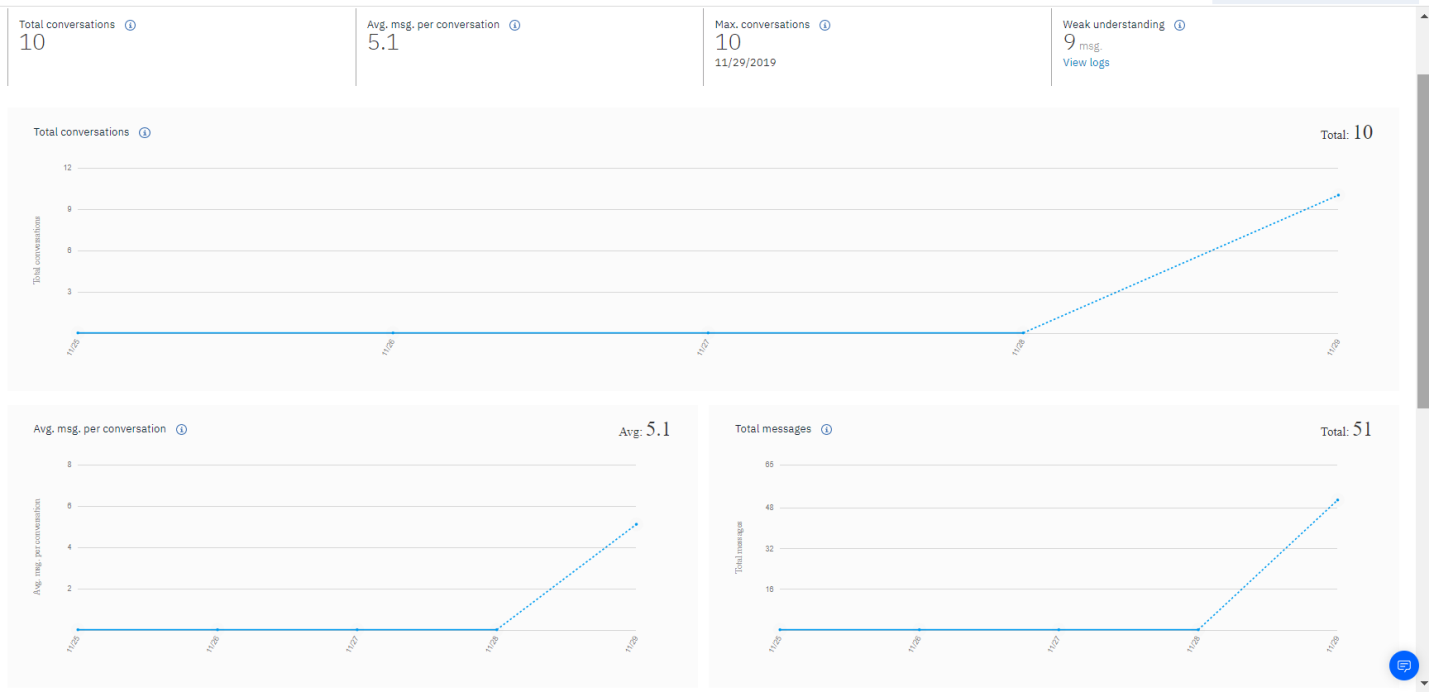
Figure 9. Testing and Training on Natural Language Processing

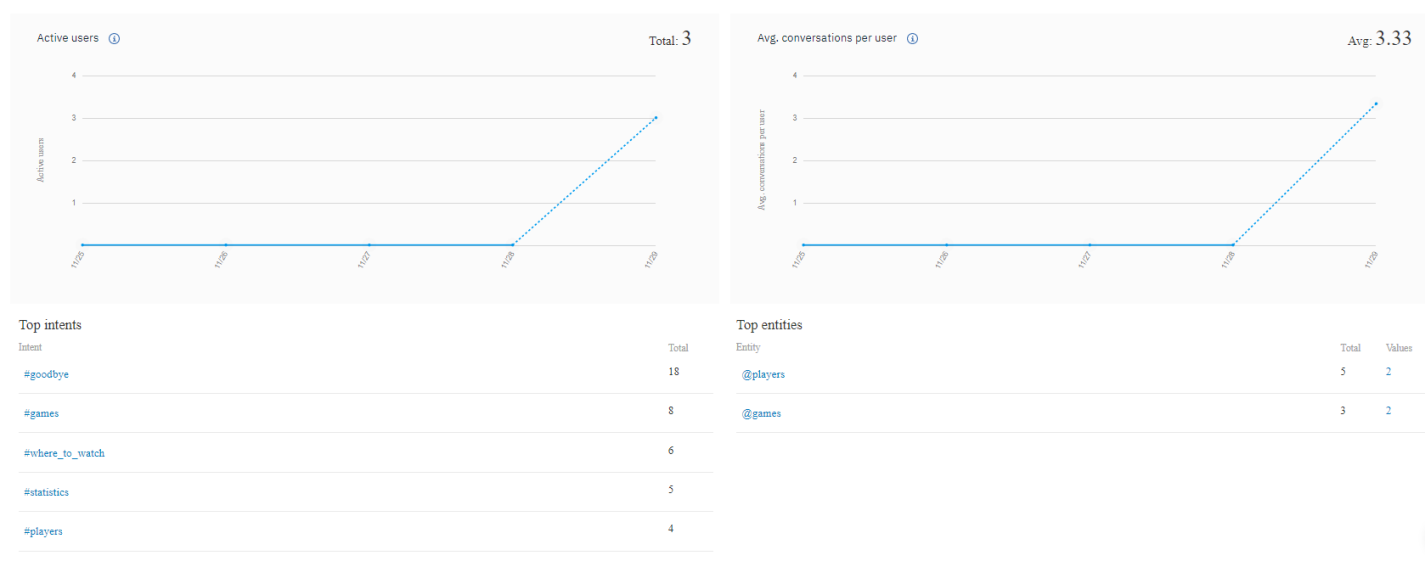


**Day 1 Experiment**

The assistant link was shared for use on November 29th, 2019[[1]](#footnote-1). A total of three users had a total of ten conversations, with an average of 5 messages per user. The biggest intents were goodbye, and games. After reviewing the logs, it appears to be because users used goodbye at a high rate no matter their purpose in using the app, and aside from that the highest desire was to know the results of the last game and when the next game was going to be held. The poorly understood input was reviewed, and it appeared that many were names inputted at the start. These were kept as irrelevant. Additionally, a user had asked about kickers which was trained to be the players intent, and a user has said ‘thanks’ which was trained to be the goodbye intent. No entities were misclassified, likely due to the well known separation of entities inherent to people familiar with the NFL. Overall, the chatbot appeared to work well, likely due to the structured nature of the experience, and the predictability of intents inherent to limiting the chatbot to only discuss a certain well understood domain, in this case the ravens football team.

Figure 10. Analytics of Experiment Day 1

****

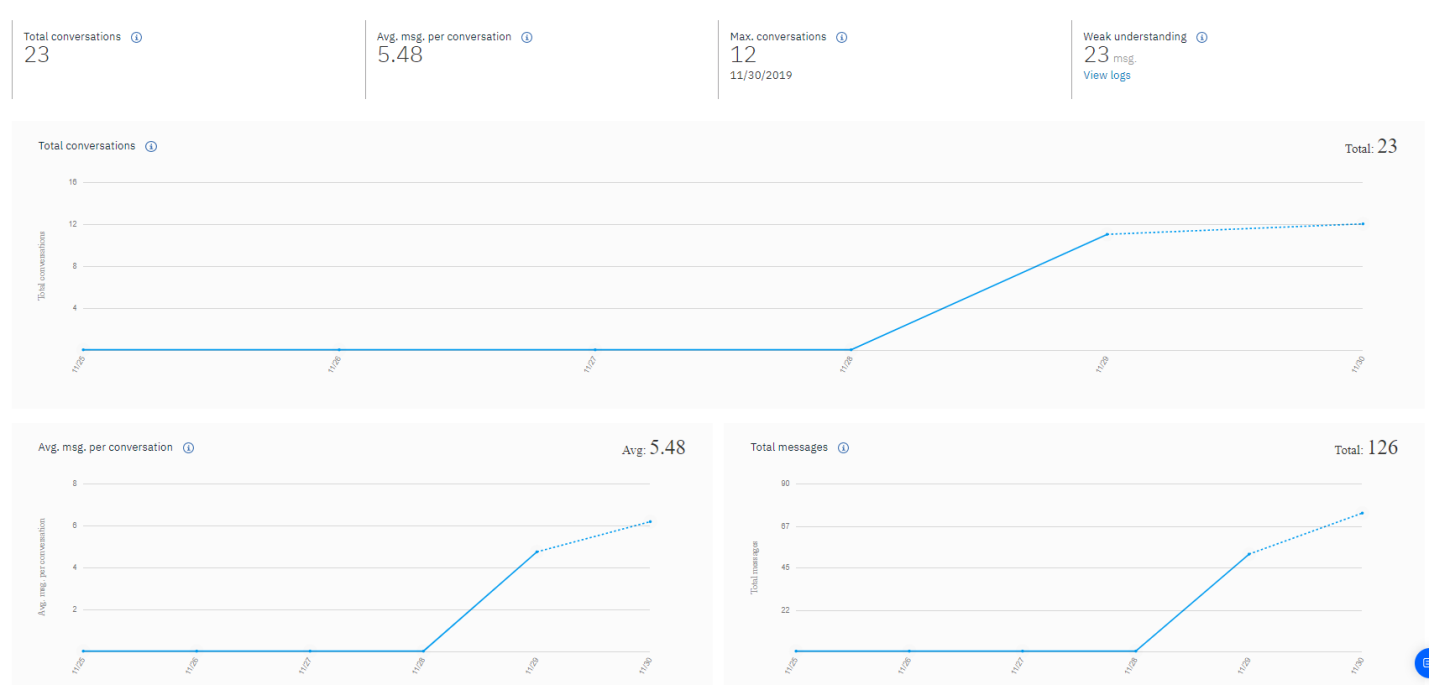
****

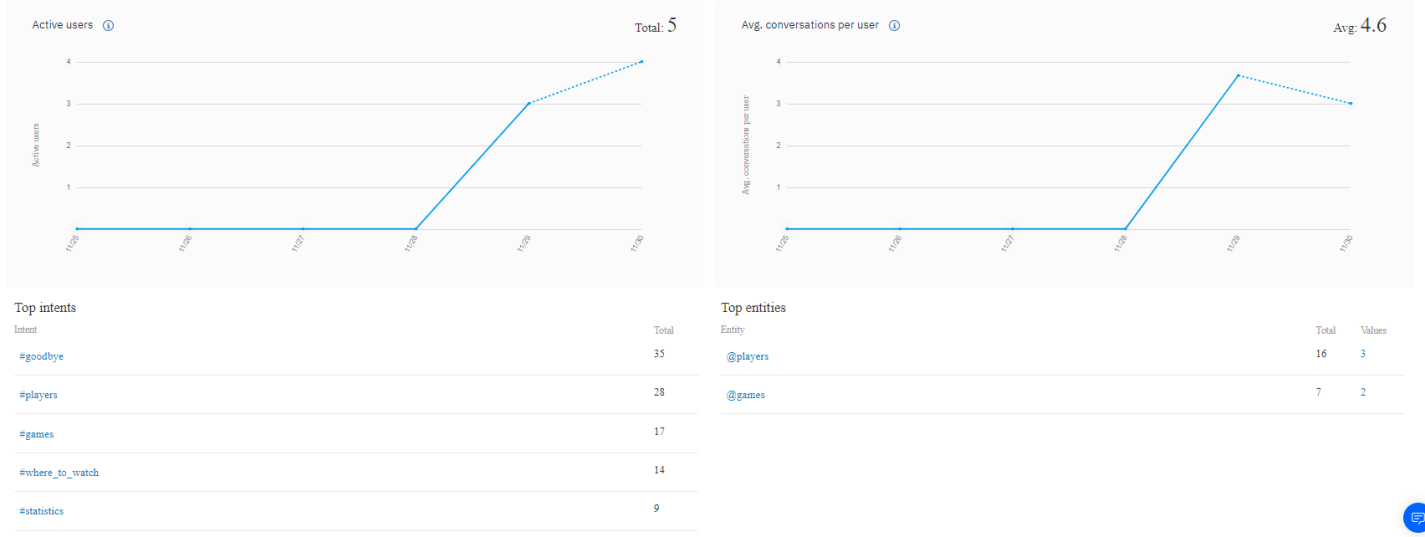
**Day 2 Experiment**

The experiment was run for another day on November 30, 2019, with roughly the same amount of conversations (**See Figure 11**). While goodbye was still the top intent, likely due to fitting into the end of any conversations, games rose in the number of intents detected. This may be due to the next football game being the next day raising interest in the subject.

When reviewing conversations, it was noted that there were more inquires regarding the kicker of the ravens and a few regarding the coach of the ravens. It was also noted that many unusual expressions for the ‘goodbye’ intent were appropriately detected. This may be because the intent is common and so has a lot of training data for the algorithm to use.

Figure 11. Experiment Day 2 Analytics

****

****

If the experiment was to be run for a third day, an additional sub-nodes of players would be introduced such as for the Ravens kicker, or even the ravens coach to see how much more demand there is for less known players. A node called “Polite” would be added for when people thank the chatbot without giving additional requests to it. The analytics would continue to reviewed to review for more synonyms for any missed intents.

**Chatbot Enhancement**

To personalize the interaction with the chatbot, images were used to give a personal connection, especially when asking about players, the dialogue was written in such a way that the information was conveyed from another person who was an enthusiastic ravens fan, and the name was retained and then reused. Additionally, the dialogue was written so that it appeared to be an actual conversation with back and forth with the user. Through these means, the desired impact was to increase engagement with the app and help reinforce the surface illusion that they were communicating with another person even when they understood it was actually a bot.

**Watson Service Integration**

One Watson service that could be integrated to improve the performance of the app would the tone analyzer. The users input could be analyzed for tone and that could be used to help understand if the bot was performing poorly and if human interaction is integrated it could be a signal to understand the user input and ask if they would like to be connected to a human.

Another Watson Service that could be beneficial could be speech to text. Many football fans are older, and dislike inputting long statements by typing especially on a phone. By allowing the user to speak what they wanted to input it could increase engagement, especially with the older crowd. Additionally, if deemed appropriate and ethically sound, you could mine the voice for information about the user such as age or gender to help target specific sections of the app which may be more relevant to them.

**Platform Deployment**

This app would be best deployed through an app, or a website. Either of these choices would be a good fit for deployment as they are both commonly used and generally accepted places where people would expect to interact with a chatbot. Additionally, these are the places people would look for information regarding Ravens. Therefore, these choices be deployed to where the demand for the information is.

**Further Training Post Deployment**

The first priority after deployment would be monitoring to see what functionality or information is being asked for by users that isn’t available so that the feature could be prioritized for the next deployment. Additionally, it would be important to ask for customer feedback and address concerns. Further, continued monitoring of conversations, especially poorly understood conversations would be essential to a successful implementation. A low skilled employee trained on the system would likely be an efficient method of monitoring and correcting the chatbot’s understanding of intents and entities rather than using the more expensive labor of a data scientist.

**Conclusion**

Chatbots are still a developing technology, although the introduction of natural language processing to help understand user input has improved the capability of IBM Virtual Assistant over previous chatbots deployed in the past. One of the challenges facing chatbots deployed today are that it takes more time to build a system that directs the user to the information that it does to simply put the information on a website and let the user find it themselves. Another challenge is that it requires constant monitoring to get the most value out of the chatbot so you can respond to changes in languages and content from week to week. Further, the more complicated the chatbot and the more intents and entities that are introduced the more difficult it will be for the chatbot to correctly pick between them.

When developing and deploying a chatbot there are best practices to help ensure the project is a success. One of the first practices it to make sure the conversations are being constantly reviewed, likely with review of a random sample as well as a review of conversations with poorly understood user input. Another best practice is to review for what people are asking the chatbot to do that it is incapable of. This way the implementation of the missing features can be made a priority. Further it is important to properly plan the dialog and nodes from the start, with planning for future implementation of new features. Finally, asking for user feedback when the goodbye node is initiated can be a method to get feedback on how it is performing.

This chatbot would not replace human interaction, as it is more targeted towards replacing the user searching for information on the web. However, human interaction could be added to it to increase user engagement and monetization. Specifically, it could direct users to human agents to help them sign up for services to watch the game, buy tickets, or even buy more merchandise. An important use case could be a human agent to help users navigate streaming services as well.

1. <https://assistant-chat-us-south.watsonplatform.net/web/public/31d58c7c-f43e-4c11-8a69-4e3d4c594fc1> [↑](#footnote-ref-1)