Web-based visualization and corpus management for language modeling

Weekly Progress Report #9

Customer
Prof. Dr. Dietrich Klakow, Spoken Language Systems Group, Saarland University

Developer
Tiyash Basu, Bhavesh Bhansali, Supratim Das, Satabdi Ganguly, Clayton Greenberg, and Rohit Kumar Gupta

Tutor
Rafaella Antonyan

Submission Date
22/01/2016

Project description:
The high level goal of this project is to create a functional, web-based graphical user inter-face (GUI) for the language modelling toolkit developed by the Spoken Language Systems Group, Saarland University (LSVLM). When possible, the design of this GUI must be close to that of the existing nonfunctional GUI in order to provide a consistent user experience for existing users. The GUI must extend the functionality of the LSVLM toolkit and existing GUI in the areas of corpus management and visualization of provided outputs (Experimentation). The intended users of this GUI are expected to know the basics of language modelling, but not command-line tools.

Progress

This week was the first week of the third iteration. We had a first meeting with dedicated beta testers on Wednesday. In the meeting, we presented all the implemented must have features. As got feedback from the users, they are satisfied with system. We got feedback for some changes for better look and minor modification in some functionality. We had an internal planning session to handle user's feedback. We have started working as per the feedback given by the users. However, we have prioritized the may have features and started working on it concurrently. We have implemented some interpolation functions in LM application. The updated code has been pushed to Git.

What has been done

- LDAP authentication

- Updated Requirements

- All Use Cases

- Class Diagram

- All Sequence Diagrams

- All Paper Prototypes

- 91% Must have features

- Prioritizing may have features

- Implemented some may have features

Ongoing work

- Modification as per the user's feedback

- The language modelling design page is in progress but not complete

- Implementation of may have features

- Scrolling on the surprisal graph is in progress but not complete

- Further visualizations would be good

Plans for the next week

We plan to work on the last 3 remaining must have features and some may have features.

Problems

Using RZ for authentication, while originally desired by the customer, requires too many outstanding dependencies, so the customer is happy to use his own LDAP directory instead. If this becomes different in the future, we will change the code then.