EE / CprE / SE 492 – Sddec18-16

Use machine learning to predict relevant support content based on historical user interactions.

Week 6 Report
November 6, 2018 – November 20, 2018

Team members
Erin Elsbernd: Machine Learning Lead
Ram Luitel: Project Manager and Software Architect
Faizul Jasmi: Communication coordinator
Taizhong Huang: QA Lead
Christian Chiang: Cloud Tech Lead
Khoa Bui: Webmaster and DB lead

Summary of Progress for the Past Two Weeks.

Solved the data reading problem that we mentioned in previous report. However, currently, it only reads one third of of data from DynamoDB table. Also implementation of reading database code on AWS is done.

Pending Issues

We would still need to do a full automation of the application process. This means fully integrate the data cleaning and the model predictions in AWS. This is beyond the scope of the project; our client (Workiva) just stated that our models should produce a 70% accuracy, but our best model produces only 64%.

Plans for Upcoming Reporting Period

The rest of the semester we will work on polishing models for use in lambdas and working on final presentation. We also need some integration work left which we think will be finished this week.

Individual Contributions

Team Member	Contribution	Biweekly hours	Total hours
Erin Elsbernd	Finalized grouping strategies for our model.Our best model gets around 64%, but with new grouping strategy we have better precision and recall for all classes. Experimented with different over-sampling techniques and PCA to reduce matrix size. Contributed to final report.	6	25
Ram Luitel	This last two week I focus more on documentation and requirements. I managed and lead the team to write poster and final report	5	24
Faizul Jasmi	Worked with Khoa Bui in tuning the Lambda Function to properly store the data in DynamoDB	2	21
Taizhong Huang	Contributed to final report. Reviewed other member's code.	1	19
Christian Chiang	Finished code for data cleaning in AWS lambdas, reviewed tables for the cleaned data and verified that everything was working properly. Contributed to the documentation and paperwork of the project.	3	24

Khoa Bui	Try to solve reading data problem. Currently, it only reads ⅓ of data from DynamoDB table. Research how to read JSON data to pandas dataset. Try to implement reading database code on AWS	6	23