

EE / CprE / SE 492 – Sddec18-16

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

Week 3 Report

September 24, 2018– October 6, 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 the Past Two Weeks.

- Exploring new data - Erin:
 - Tested new data with old data cleaning script.
 - Combined new data with old and ran through classification models.
 - More data lead to a slight bump in overall accuracy of model.
- Data cleaning and feature extraction - Ram:
 - Cleaning new data that our client provide last week so that I can used it with my current model and try more machine learning models.
 - My cleaning script is not good as a result fetching it to my current models does give any better result.
- Review code - Taizhong:
 - Read new data.
 - Research new algorithm and neural network.
- Receiving and collecting new data from Workiva - Faizul:
 - Actively in contact with Workiva to acquire the datasets that we need.
 - Made sure that the data received is the same format as before.
 - Distributed the data to the rest of the team for further processing.
- Review AWS requirement - Christian:
 - Connection between components complete
 - Microservice basic template done

- Tested connectivity and permissions between account
- Project website update - Khoa:
 - Update the website with all current item. Now we have all report for this semester as well.
 - Try to clean and new data that client provide last week.

Pending Issues

Our pending issue is same as previous as our model's performance is not where it needs to be in terms of article-group prediction accuracy. We are hoping that using this more data, that our client provide us last week our models will be able to perform at a higher level. Our current block is we need to clean and build more machine learning model with new data.

Plans for Upcoming Reporting Period

- Article Grouping with new data- Erin:
 - Continue work on having the model output a ranked confidence of predictions, e.g. 85% Group 1, 10% Group, 5% Group 3 etc.
 - Now that there is finally more data, continue with experimenting with creating optimal groups for random forest classification.
- New data cleaning - Ram:
 - Finally our client provided more data. I am trying to clean this data so that I can used it with my current model and try more machine learning models.
 - I am also trying to extract different features so that I can apply different machine learning model to come up with better prediction.
- Faizul:
 - More data, the better. I will keep on updating with Workiva and ask them to feed us with more data. Having a large set of it really helps in training our models and see their performances to further analyze which one is better.
 - I will also make sure that the data is distributed correctly among the team members.
 - Continue working on porting Jupyter files to .py
- New Features - Taizhong:
 - I will try to build a new neural network model with our new data.

- Christian:
 - Incorporate scripts into the microservices
 - Improve performance of scrips to adapt to limitations
 - Figure out a way to speed up reading and writing of the models
- Update project website - Khoa:
 - Keep updating website and contribute and collaborate with team member for any database needs for storing our data and deploying our models.

Individual Contributions

Team Member	Contribution	Biweekly hours	Total hours
Erin Elsbernd	Explored any new insights with additional data and got increased model accuracy.	2	10
Ram Luitel	Meet with Christian to discuss possible solutions for lamda processing issue and try to clean new data to extract feature to feed my current models.	2	10
Faizul Jasmi	Finishing porting .ipynb file to .py Received & collected new data and distributed among team members	2	11
Taizhong Huang	Research new algorithm and neural network.	2	7

Christian Chiang	Connection between components complete Microservice basic template done Tested connectivity and permissions between account	4	15
Khoa Bui	Meet with Christian to discuss possible solutions for lamda processing issue and update website.	2	8