sddec18-16: Use machine learning to predict relevant support content based on historical user
Week 7 Report
March 4 - March 17

Team Members

Erin Elsbernd — *Communication Coordinator and Machine Learning Lead*
Ram Luitel — *Project Manager & Software Architect*
Faizul Jasmi — *Testing & AWS Tech Lead*
Khoa Bui — *Database & Web Master #2*
Taizhong Huang — *Testing*
Christian Chiang — *Web Master & AWS Tech Lead*

Summary of Progress this Report

This two week we really focus on feature extracting and cleaning the raw data that we received from our client and started processing it into viable training data to use. We also began selecting models to use for the given data and implementing them. We have broken into 2 sub-teams, where one team will focus on neural networks, and the other on trees and recommendation systems. We came up with initial results which predicted roughly 14% accurately but after meeting with faculty advisor we have the more clear idea to improve the performance.

Pending Issues

We still need to do more work on clean our initial data which consist of more than 1000 cvs file. We also need to play more with our clean by extracting different features and try out more model to compare the results.

Plans for Upcoming Reporting Period

This Thursday we will be meeting with our client for an update on our progress and we have some question regarding data. We will ask for clarification and keep working on our models and data by extracting more features to meet our client need which is at least 70% accurately prediction.

Individual Contributions

Team Member	Contribution	Weekly Hours	Total Hours
Erin Elsbernd	Finished processing the Workiva data. Mapped the action events and titles to int values. Started using the data to fit a neural network with LSTM	10	53
Ram Luitel	Try to link the aggregated data that Erin cleans up with the end data that client provided. I am trying to use collaborative filtering and trying to recommend article according to behavior, so I am in the process of it. I also try to use. I have also I have tried Naive Bayes classification model using title and issues to classify which article would be	8	42

	helpful and which are not.		
Faizul Jasmi	Research on LSTM, revised project plan timeline and project risks and expenditures and research on AWS to further prepare implementation phase	5	37
Khoa Bui	Update the website by adding documents and weekly reports. Play with data to understand random forest model and how does it really work. Also learned some python library that I could apply to build the predicting model. I also help to revised project plan version two especially focus on re-building two-semester timeline, testing plan, and project process.	7	39
Taizhong Huang	Helped to contribute design document. Tried to clean data and try random forest model.	7	33
Christian Chiang	I did research about AWS and how hosting and microservices worked there. I also try to test some hosting on lambda and EC2 instances. Also read about hosting different type of models and which service would work best for each of them.	5	37