sddec18-16: Use machine learning to predict relevant support content based on historical user
Week 5 Report
February 17 - February 24

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

As we received more data, we spent most of our time in data exploration and processing. We have also started building random forest and neural network models based on the research findings that we got from previous weeks. We have also made feature generation scripts.

Pending Issues

We received data that is spread out into over 40 csv files, which some of them contains different labels that are still related. Thus, the issue is to find the relations between the labels to make it as effective data points. Some correct relations are even from with a file that we received weeks ago.

Plans for Upcoming Reporting Period

We will meet again with Alex to go thru the data. In the meantime, we will work on labeling the data ourselves based on what we have already understood.

Individual Contributions

Team Member	Contribution	Weekly Hours	Total Hours
Erin Elsbernd	Researched time series models like ARIMA and implemented it with sample data. Worked with time series data. Researched on RNN models. Implemented basic CNN. Implemented term frequency-doc frequency (tf-idf) matrix with new data. Contributed for Design Document.	8	36
Ram Luitel	Made power points for lightning talk two. Wrote the logistic regression model for data. Worked on random forest model. Wrote a practice sample random forest algorithm. Contributed on Design Document.	6	30
Faizul Jasmi	Explored the data and seek for possible	5	27

	relations between the files. Researched on possible ways to process the data points to make it as effective inputs-JavaScript w/ jQuery or Panda w/ Python.		
Khoa Bui	Fixed details on group website and restructured the code. Drew diagram for lightning talk and redo project timeline. Contributed on Design Document	7	28
Taizhong Huang	Setting up the environment and installed libraries needed. Contributed on Design Document.	3	22
Christian Chiang	Researched on different types of RNN specifically on how they would work with the data from WDesk. Looked for alternative for LSTM. Contributed on Design Document.	6	28